ERIC Document Resume	See Instruction Sheet	DATE OF RESUME	DATE RECEIVED

A. PUBLICATION DATA

1. ERIC NUMBER: ED002569
2. ORIGINATOR CONTROL NO(S).: N.A.
3. IDENTIFICATION NO.: CR 1391
4. AUTHOR(S): Lambert, Philip and others
5. TITLE OF DOCUMENT ABSTRACTED: Classroom Interaction, Pupil Achievement and Adjustment in Team Teaching as Compared with the Self-Contained Classroom
6. PUBLICATION TITLE: N.A.
7. EDITOR(S): N.A.
8. PLACE OF ORIGIN
ORG. University of Wisconsin
ST. Madison, Wisconsin
CITY, STATE
TELE. EXT.
INST./LOC CODE
9. PUBLISHER OR SOURCE
University of Wisconsin
Madison, Wisconsin

10. VOL.	11. PAGES	12. MO./YR.	13. TOTAL PP.
N.A.	N.A.	1964	272

DO NOT WRITE IN SPACE BELOW

MICROFILM COPY AVAILABLE
YES ☒ NO ☐

COPYRIGHT GRANTED
YES ☐ NO ☒

COPYRIGHT RELEASE FILED
YES ☐ NO ☐

TYPE OF REPORT
PROPOSAL ☐
IN-PROGRESS ☐
FINAL ☒
FOLLOW-UP ☐

OTHER ERIC REFERENCES:

ERIC SATELLITE CODE

B. INDEXING DATA

14. SUBJECT AREA
15. TYPE OF ENDEAVOR: Research

16. INDEXING TERMS (6-12)	student achievement	classroom interactions	
teacher training	student personality	self-contained classroom	
team teaching	student discipline	student adjustment	
elementary education	classroom methods		

17. ABSTRACT (MAX. 250 WORDS)

Explored are teaching and learning relationships or forms of classroom interaction in the team teaching arrangement, and the differences between these and traditional classroom interaction with respect to student adjustment and student achievement. An attempt was made to improve the sophistication of educational research instruments. A two-year team teaching program involving an entire elementary school formed the basis of study. Conclusions were that significant differences in classroom interaction and in student achievement exist between the team and self-contained organizations; but no significant differences were found in personal and social adjustment, in teacher awareness of student characteristics, in absenteeism, in the frequency of discipline infractions, and in changes in the social structure within the classes organized on a team-teaching basis and those organized on a self-contained basis. Major emphasis was placed on the methodology of research and the analysis of data. Extensive tables are included.

ERIC No.
18. SOURCE OF ABSTRACT

C. FUNDING DATA

19. SUPPORTING ORGANIZATION
20. DIVISION OR DEPT.
21. CONTRACT NO.
PROJECT NO.
22. TOTAL COST OF PROJECT
23. DURATION OF PROJECT

U.S. Department of Health, Education, and Welfare
Office of Education

ERIC Form 1

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education

This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

ED002569

CLASSROOM INTERACTION, PUPIL ACHIEVEMENT AND ADJUSTMENT IN TEAM TEACHING AS COMPARED WITH THE SELF-CONTAINED CLASSROOM

Cooperative Research Project No. 1391

Philip Lambert
Director

William Wiersma
William L. Goodwin
Richard F. Roberts
Co-Authors

University of Wisconsin
Madison, Wisconsin
1964

The research reported herein was supported by the Cooperative Research Program of the Office of Education, U. S. Department of Health, Education, and Welfare.

PRECEDING PAGE BLANK– NOT FILMED

ACKNOWLEDGMENTS

We are highly indebted to the following for their various contributions towards the present research:

Teachers and students in the two cooperating elementary schools in Madison, Wisconsin, and particularly their Principals, Mrs. Helen Bruce of Washington School, and Miss Helen Simon of Longfellow School.

Superintendent Robert D. Gilberts and former Superintendent Philip H. Falk of the Madison Public Schools without whose strong backing this project would not have been possible.

Professor John Guy Fowlkes, Director of the Wisconsin Improvement Program, for jointly sponsoring a pilot study upon which much of this research was based.

Members of the research staff including: David R. Stucki, Michael A. Rieck, and David E. Wiley, each of whom was a research assistant for a period of time; Mrs. Elinor A. Elfner who conducted the side study on Peer Structure; and Mrs. Dorothy Hougum who served as secretary for the Project.

Professor Ned A. Flanders of the University of Michigan, and Professor John Withall at that time of the University of Wisconsin for their advice on student-teacher interaction models; Professor Julian C. Stanley, Director of the University of Wisconsin Laboratory for Experimental Design, for advice on the design of the project; and Professor Edgar F. Borgatta, Chairman of the Sociology Department, University of Wisconsin.

PRECEDING PAGE BLANK- NOT FILMED

PREFATORY NOTE

Shortly before this report was printed in final form the first full-length book on team teaching was released. * The brief interim between the book's release and the publication of this report precluded weaving its vast contents into the review of the literature. At the same time, it was deemed advisable to mention the book here to alert the reader to this pioneering publication in the area of team teaching. Seven authors (Judson T. Shaplin, Henry F. Olds, Jr., Robert H. Anderson, Joseph C. Grannis, Cyril G. Sargent, Dan C. Lortie, and Glen Heathers) have each written one or more of the eleven different chapters in the book. The Appendix includes an extensive bibliography as well as a listing of many team teaching projects.

A large portion of the material in the book is applicable to this project. Especially pertinent are passages dealing with (1) the Wisconsin Improvement Program and interns on the teams (pp. 53, 174, and 201-202); and (2) research on team teaching (Chapter 10), most notably with regard to student achievement (pp. 326-331) and student adjustment (pp. 331-335). It appears necessary to point out, however, that this writer is in disagreement with Heather's implication that a one-group pre-test post-test design should often be employed in educational research. **

---

* Team Teaching, edited by Judson T. Shaplin and Henry F. Olds, Jr., New York, Evanston, and London: Harper and Row, 1964.

** In this regard, see Campbell, Donald T. and Julian C. Stanley, "Experimental and Quasi-Experimental Designs for Research on Teaching," Chapter 5 in Handbook of Research on Teaching, edited by N. L. Gage, Chicago: Rand McNally and Company, 1963.

PRECEDING PAGE BLANK- NOT FILMED

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS iii

PREFATORY NOTE v

TABLE OF CONTENTS vii

LIST OF SCHEDULES AND TABLES ix

LIST OF FIGURES xviii

PART ONE: INTRODUCTION 1

I. Problem 3
II. Objectives 5
III. Related Research: A Review of the Literature 6
IV. Statistical Methods 20
V. Populations and Experimental Design 30

PART TWO: CLASSROOM INTERACTION 37

I. Introduction 39
II. Observation 48
III. Data 55
IV. Statistical Analysis 59
V. Descriptive Analysis 85
VI. Conclusions 104

PART THREE: STUDENT ACHIEVEMENT AND PERSONALITY 111

I. Achievement and Personality: Introduction 113
II. Personality: Measurement 117
III. Personality: Statistical Analysis 120
IV. Achievement: Measurement 137
V. Achievement: Statistical Analysis 143
VI. Achievement: Descriptive Analysis 170
VII. Achievement and Personality: Conclusions 176

Page

PART FOUR: SIDE STUDIES 181

I. Teacher Awareness of Student Characteristics 183
II. Distribution of Absence 191
III. Student Discipline 192
IV. Social Structure 201

PART FIVE: DISCUSSION AND CONCLUSIONS 215

I. Discussion 217
II. Conclusions 221

APPENDIX A: INTERACTION MODELS 223

I. Withall Model 225
II. Flanders Model 230

APPENDIX B: TABLES OF ADJUSTED MEANS 231

I. California Test of Personality 233
II. California Achievement Test 236

APPENDIX C: TEACHER AWARENESS 239

BIBLIOGRAPHY 245

LIST OF SCHEDULES AND TABLES

Schedule Page

1. Interaction Studies, 1961-63 43

2. Personality and Achievement, 1961-63 113

Table

1. Design of a 2x2 Factorial ANOVA of Amount of Talking by Sex and Organization 21

2. F-Ratios for Amount of Talking by Sex and Organization 21

3. Mean Percentage of Talking by Organization 25

4. Average I.Q.'s by Sex, Grade, and Organization 34

5. Distribution of 30-Minute Interaction Observations Per Teacher: First Semester, 1961-62 49

6. Distribution of Interaction Observations Per Teacher 50

7. Mean Number of Frequencies of Revised Withall Categories by Organization and Grade Level: First Semester, 1961-62 ... 60

8. Division of 2x3 Analysis of Variance for Revised Withall Categories: First Semester, 1961-62 61

9. F-Ratios of Revised Withall Categories by Category and Source: First Semester, 1961-62 62

10. Mean Number of Frequencies of Revised Withall Categories by Organization: First Semester, 1961-62 63

11. Mean Number of Frequencies of Revised Withall Categories by Grade Level: First Semester, 1961-62 64

12. Mean Percentage Scores on Revised Flanders Categories by Grade Level and Organization: First Semester, 1961-62 ... 65

Table Page

13. F-Ratios of Revised Flanders Categories by Category and Source: First Semester, 1961-62 66

14. Mean Percentage Scores of Revised Flanders Categories by Organization: First Semester, 1961-62 67

15. Mean Percentage Scores of Revised Flanders Categories by Grade Level: First Semester, 1961-62 68

16. Mean Percentages of Revised Flanders Categories for the Team by Type of Teacher and Grade Level: First Semester, 1961-62 69

17. Division of 2x2 Analysis of Variance for Revised Flanders Categories: First Semester, 1961-62 70

18. F-Ratios of Revised Flanders Categories for Team Teachers by Category and Source: First Semester, 1961-62 70

19. Mean Percentages of Revised Flanders Categories for the Team by Grade Level: First Semester 1961-62.................. 71

20. Mean Percentages of Revised Flanders Categories for the Team by Type of Teacher: First Semester, 1961-62 72

21. Mean Number of Frequencies of Revised Withall Categories for Arithmetic by Grade Level: Second Semester, 1961-62 73

22. Mean Number of Frequencies of Revised Withall Categories for Science by Grade Level: Second Semester, 1961-62 73

23. Mean Number of Frequencies of Revised Withall Categories for Social Studies by Grade Level: Second Semester, 1961-62 . 74

24. Mean Number of Frequencies of Revised Withall Categories for Reading by Grade Level: Second Semester, 1961-62 74

25. Division of 4x6 Analysis of Variance for Revised Withall Categories: Second Semester, 1961-62 75

26. F-Ratios of Revised Withall Categories for Self-Contained Classes by Category and Source: Second Semester, 1961-62 .. 76

Table Page

27. Mean Number of Frequencies of Revised Withall Categories by Subject Area and Category: Second Semester, 1961-62..... 77

28. Mean Number of Frequencies of Revised Withall Category 3 by Subject Area and Grade Level: Second Semester, 1961-62 . 78

29. Mean Number of Frequencies of Flanders Categories for Arithmetic by Organization: First Semester, 1962-63 79

30. Mean Number of Frequencies of Flanders Categories for Social Studies by Organization: First Semester, 1962-63 80

31. Mean Number of Frequencies of Flanders Categories for Reading by Organization: First Semester, 1962-63 81

32. Division of 3x3 Analysis of Variance for Flanders Categories: First Semester, 1962-63........................ 82

33. F-Ratios of Flanders Categories by Category and Source: First Semester, 1962-63 82

34. Mean Number of Frequencies of Flanders Categories by Organization: First Semester, 1962-63 83

35. Mean Number of Frequencies of Flanders Category 3 by Organization and Subject Area: First Semester, 1962-63 84

36. Interaction Areas of the Flanders Matrix 86

37. Percentages of Total Observations of Flanders Interaction Areas by Organization: First Semester, 1962-63 87

38. Percentages of Total Observations of Flanders Interaction Areas by Organization and Grade Level: First Semester, 1962-63 89

39. Percentages of Selected Observations of Flanders Interaction Areas by Type of Teacher and Grade Level: First Semester, 1962-63 102

40. Summary of Significantly Different Interaction Categories by Organization: 1961-63....................... 105

Table Page

41. Summary of Significantly Different Interaction Categories by Grade Level: First Semester, 1961-62 106

42. Mean Percentages of Flanders Categories for the Team by Type of Teacher, Grade Level, and Category Comparing the Studies of First Semester 1961-62 and the First Semester 1962-63.. 108

43. Personality Test Level by Grade and Date: 1961-63 117

44. Distribution of Students Participating in Personality Tests by Grade Sex and Organization: Academic Year 1961-62 118

45. Distribution of Students Participating in Personality Tests by Grade, Sex and Organization: Academic Year 1962-63 119

46. California Test of Personality: Personal Adjustment Pre- and Post-Test Mean Scores by Sex, Grade and Organization: Academic Year 1961-62 120

47. California Test of Personality: Social Adjustment Pre- and Post- Test Mean Scores by Sex, Grade and Organization: Academic Year 1961-62 121

48. California Test of Personality: Total Adjustment Pre- and Post-Test Mean Scores by Sex, Grade and Organization: Academic Year 1961-62 122

49. F-Ratios of California Test of Personality Scores by Type of Adjustment and Source: Academic Year 1961-62 123

50. California Test of Personality: Personal Adjustment Overall Mean Scores by Organization: Academic Year 1961-62......... 124

51. California Test of Personality Mean Scores by Sex, Grade and Organization; Primary Level: Academic Year 1962-63 133

52. California Test of Personality Mean Scores by Sex, Grade and Organization; Intermediate Level: Academic Year 1962-63 . . 134

53. F-Ratios of California Test of Personality Scores on the Primary Level by Type of Adjustment and Source: Academic Year 1962-63 135

Table | Page

54. F-Ratios of California Test of Personality Scores on the Intermediate Level by Type of Adjustment and Source: Academic Year 1962-63 136

55. Achievement Test by Grade, Test Level and Date 137

56. Achievement Test by Grade, Test Level and Date: May, 1963 138

57. Range of Reliability Coefficients on the Different Levels of the California Achievement Tests 139

58. Iowa Test of Basic Skills Reliability Coefficients by Subject Area and Grade 140

59. Distribution of Students Participating in Achievement Tests by Grade, Sex and Organization: Academic Year 1961-62 141

60. Distribution of Students Participating in Achievement Tests by Grade, Sex and Organization: Academic Year 1962-63 142

61. California Achievement Test: Reading Mean Grade Placement, Pre- and Post- Tests by Sex, Grade and Organization; Grades 2-6; Academic Year 1961-62 143

62. California Achievement Test: Arithmetic Mean Grade Placement, Pre- and Post- Tests by Sex, Grade and Organization; Grades 2-6: Academic Year 1961-62 145

63. California Achievement Test: Language Mean Grade Placement, Pre- and Post- Tests by Sex, Grade and Organization; Grades 2-6: Academic Year 1961-62 145

64. California Achievement Test: Total Achievement Mean Grade Placement, Pre- and Post- Tests by Sex, Grade and Organization; Grades 2-6: Academic Year 1961-62 146

65. F-Ratios of California Achievement Test Scores by Subject Area and Source: Academic Year 1961-62 147

66. California Achievement Test: Overall Raw Mean Grade Placement by Organization for Reading, Arithmetic, Language, and Total Scores: Academic Year 1961-62 148

Table Page

67. California Achievement Test: Arithmetic Adjusted Means by Sex and Grade: Academic Year 1961-62 155

68. California Achievement Test: Reading, Arithmetic, Language and Total Achievement Mean Grade Placements by Sex, Grade, and Organization; First Grade: Academic Year 1962-1963 156

69. F-Ratios of California Achievement Test Scores for First Grade by Subject Area and Source: Academic Year 1962-63 ... 157

70. California Achievement Test: Reading, Language, and Total Achievement Overall Mean Grade Placements by Organization; First Grade: Academic Year 1962-63 157

71. California Achievement Test: Language and Total Achievement Overall Mean Grade Placements by Sex; First Grade: Academic Year 1962-63 158

72. California Achievement Test: Reading, Arithmetic, Language, and Total Achievement Mean Grade Placements by Sex, Grade, and Organization; Second Grade: Academic Year 1962-63..... 159

73. F-Ratios of California Achievement Test Scores for Second Grade by Subject Area and Source: Academic Year 1962-63.. 159

74. California Achievement Test: Reading, Language, and Total Achievement Overall Mean Grade Placements by Organization; Second Grade: Academic Year 1962-63 160

75. Iowa Test of Basic Skills: Vocabulary Mean Grade Placements by Sex, Grade and Organization; Third-Sixth Grades: Academic Year 1962-63 161

76. Iowa Test of Basic Skills: Reading Mean Grade Placement by Sex, Grade and Organization; Third-Sixth Grades; Academic Year 1962-63 162

77. Iowa Test of Basic Skills: Total Language Mean Grade Placement by Sex, Grade and Organization; Third-Sixth Grades: Academic Year 1962-63 163

Table Page

78. Iowa Test of Basic Skills: Total Work Study Skills Mean Grade Placement by Sex, Grade, and Organization; Third-Sixth Grades: Academic Year 1962-63 164

79. Iowa Test of Basic Skills: Total Arithmetic Mean Grade Placement by Sex, Grade and Organization; Third-Sixth Grades: Academic Year 1962-63 165

80. Iowa Test of Basic Skills: Composite Mean Grade Placement by Sex, Grade and Organization; Third-Sixth Grades: Academic Year 1962-63 166

81. F-Ratios of Iowa Test of Basic Skills Scores for Third-Sixth Grades by Subject Area and Source: Academic Year 1962-63 .. 167

82. Iowa Test of Basic Skills: Overall Mean Grade Placement by Subject and Organization; Third-Sixth Grades: Academic Year 1962-63 168

83. Iowa Test of Basic Skills: Total Language Mean Grade Placement by Grade and Sex; Third-Sixth Grades: Academic Year 1962-63 169

84. Iowa Test of Basic Skills: Composite Mean Grade Placement by Grade and Sex; Third-Sixth Grades: Academic Year 1962-63 169

85. Tests Used in Descriptive Analysis of Student Achievement: Academic Years 1961-63 171

86. Relative Differences in Mean Grade Placement Between Organizations, on Achievement Scores: Academic Years 1961-63 172

87. F-Ratios of Teacher Awareness Responses (Questions 1, 2, and 4-7) for First and Second Administrations by Source: Academic Year 1961-62 187

88. Percents of Correct to Total Responses on Question 3 of the Teacher Awareness Questionnaire for First and Second Administrations by Organization and Level: Academic Year 1961-62 188

89. Percents of Correct to Total Responses on Question 3 of the Teacher Awareness Questionnaire for First Administration by Type of Team Teacher and Level: Academic Year 1961-62 ... 189

Table Page

90. F-Ratios of Teacher Awareness Responses (Question 8) for First Administration by Source: Academic Year 1961-62 189

91. Mean Number of Half-Days Absent by Organization and Instructional Level: Academic Year 1961-62 191

92. H Values of the Kruskal-Wallis Test by Category and Comparison Group 198

93. Frequencies and Chi-Square Values by Level, Group, and Category 199

94. Distribution of Students for Social Structure Study by Organization and Grade 202

95. Summary of the Procedure of Part I of the Analysis Based on Chi-Square Tests 205

96. Fifteen Transition Categories Indicating Direction and Amount of Change in Relationships between Two Points in Time 206

97. Summary of the Procedure of Part II of the Analysis Based on Chi-Square Tests 207

98. Chi-Square Values for Order of Dependence of Markov Chain in Longfellow School, Sixth Grade 208

99. Chi-Square Values Obtained from Comparison of Transition Probabilities of Self-Contained Classrooms with Team-Teaching Organization 209

100. Chi-Square Values Obtained from Comparison of Transition Probabilities of Self-Contained Classrooms within Schools .. 211

101. Chi-Square Values Obtained from Comparison of Distributions of Types of Transitions in Self-Contained Classrooms 212

102. Chi-Square Values Obtained from Comparison of Distributions of Types of Transitions of Team-Teaching Organization with Each Self-Contained Classroom 213

Table | Page

103. California Test of Personality: Personal Adjustment Adjusted Means by Sex, Grade, and Organization: Academic Year 1961-62 233

104. California Test of Personality: Social Adjustment Adjusted Means by Sex, Grade, and Organization: Academic Year 1961-62 234

105. California Test of Personality: Total Adjustment Adjusted Means by Sex, Grade, and Organization: Academic Year 1961-62 235

106. California Achievement Test: Reading Adjusted Means by Grade and Organization: Academic Year 1961-62 236

107. California Achievement Test: Arithmetic Adjusted Means by Grade and Organization: Academic Year 1961-62 236

108. California Achievement Test: Language Adjusted Means by Grade and Organization: Academic Year 1961-62 237

109. California Achievement Test: Total Achievement Adjusted Means by Grade and Organization: Academic Year 1961-62 .. 237

LIST OF FIGURES

Figure Page

1. Amount of talking by sex and organization 23
2. Frequency of number of sentences by sex 23
3. Flanders matrix, showing tabulation of a typical twelve-second classroom interaction 42
4. Indirect teacher influence by organization and grade level . 91
5. Direct teacher influence by organization and grade level . . 92
6. Student talk by organization and grade level.............. 93
7. Silence or confusion and student discipline by organization and grade level 94
8. Content cross by organization and grade level 95
9. Indirect teacher influence by type of teacher and grade level 97
10. Direct teacher influence by type of teacher and grade level. 98
11. Student talk by type of teacher and grade level 99
12. Silence or confusion and student discipline by type of teacher and grade level 100
13. Content cross by type of teacher and grade level 101
14. California Test of Personality: Adjusted means for personal adjustment by sex and grade 126
15. California Test of Personality: Adjusted means for social adjustment by sex and grade 127
16. California Test of Personality: Adjusted means for total adjustment by sex and grade 128
17. California Test of Personality: Adjusted means for personal adjustment by organization and grade 129

Figure Page

18. California Test of Personality: Adjusted means for social adjustment by organization and grade 130

19. California Test of Personality: Adjusted means for total adjustment by organization and grade 131

20. California Achievement Test: Adjusted means for reading by grade and organization 150

21. California Achievement Test: Adjusted means for arithmetic by grade and organization 151

22. California Achievement Test: Adjusted means for language by grade and organization 152

23. California Achievement Test: Adjusted means for total achievement by grade and organization............... 153

24. California Achievement Test: Net gain in grade placement for total achievement by grade and organization 154

25. California Achievement Test: Mean grade placements for total achievement post-test by grade and organization; 1961-1962 178

26. California Achievement Test and Iowa Test of Basic Skills: Mean grade placements for total achievement (CAT) and composite (ITBS) post-tests by grade and organization; 1962-1963 179

PRECEDING PAGES BLANK—NOT FILMED

I. PROBLEM

Among current developments in education, there is one which is causing interest and excitement: team teaching. In its simplest form, team teaching occurs whenever more than one teacher has responsibility for the same group of children at a given time -- for example, whenever a certified teacher is assisted by a student teacher. In a more complex form the teaching team is a group of teachers of differing experience and ability aided by various non-professionals. The team may have from three to six, or more, members and be responsible for a class of as many as 150 children. The children may be all of the same grade level or of different grade levels. Team teaching should not be confused with the departmentalized plan of classroom organization in which each teacher is responsible for one subject area. The members of an instructional team, though they are each specially competent in certain areas, have a primary responsibility toward all the children in the large class. For example, when one teacher presents a science demonstration to the entire class, the other teachers assist her during the presentation. When the class breaks up into small work groups, all the teachers have heard and participated in the demonstration. This enables them to be effective small-group leaders.

This report concerns the kinds of teaching and learning relationships, or forms of classroom interaction in the team teaching arrangement, and the differences between these and traditional classroom interaction. A related question studied is student adjustment and student achievement in the team structure as compared to the self-contained classroom. Differences or similarities in interaction might be related to differences (or similarities) in adjustment and achievement.

A two-year term teaching program involving an entire elementary school formed the basis of this study. One additional organization which used the traditional self-contained classroom approach was studied in parallel. Over the two-year period, classroom interaction was studied in several different ways, using different models of student-teacher interaction. Student achievement and personality measures were also varied and subjected to different methods of analysis.

Major emphasis in this report has been placed on the methodology of research and the analysis of data. Even comparatively simple questions about teaching and learning processes can necessitate carefully controlled experimentation and complicated statistical analyses. Many of the experimental methods, therefore, which are presented in this report, may be unfamiliar to the educational administrator. For this reason care has been taken to explain the research and statistics in greater detail than is usually undertaken in a technical report. In particular, a section explaining the fundamental principles of the statistical methods has been included in the Introduction for the sake of readers who are unacquainted with the specific techniques.* (Readers who have statistical training, on the contrary, will find it unnecessary to read this section.)

It is hoped, finally, that the present report will make not only a contribution to our knowledge of team teaching, but also it will add to our understanding of the methods of research and analysis in the educational field.

---

*See Part IV: Statistical Methods, page 20 below.

II. OBJECTIVES

A. General Hypotheses

The three general hypotheses tested in this experiment were:

1. There is no significant difference in classroom interaction between the self-contained and team organizations.

2. There is no significant difference in student adjustment between the self-contained and team organizations.

3. There is no significant difference in school achievement between the self-contained and team organizations.

B. Related Hypotheses

In addition to the three primary hypotheses stated above, the following hypotheses were tested in related studies:

1. There is no significant difference in teacher awareness of student characteristics between the teachers of the team and self-contained organizations.

2. There is no significant difference in absenteeism between the students of the team and self-contained organizations.

3. There is no significant difference in frequency of discipline infractions between the classes of the team and self-contained organizations.

4. There is no significant difference in changes in the social structure within the classes between the team and self-contained organizations.

III. RELATED RESEARCH: A REVIEW OF THE LITERATURE

The numerous articles on team teaching in the past few years necessitate an extensive review of the literature in order to put this particular project in perspective. This section has been subdivided into three sections as indicated below:

A. General Aspects of Team Teaching.

B. Interaction as Related to Team Teaching.

C. Personality and Achievement as Related to Team Teaching.

A. General Aspects of Team Teaching

The general public is not critically aware of the phenomenon of team teaching. Even the well-read individual is likely to encounter widely divergent and misleading views on team teaching. For example, consider the statement, "In the elementary school, the idea of the self-contained classroom has prevailed," (Norton, 121, p. 89). It implies that team teaching has either been tried and discarded or not tried at all on the elementary level; such is not the case. Many persons may have read the following opinion on team teaching which appeared in a widely-read magazine.

> Fifty children inside four walls are too many, no matter how many adults are present. The perfectly natural hums and squeaks and rustles of first graders become a dull roar and must be hushed. In one room I watched an aide spend an entire twenty-minute period walking up and down the aisles motioning children to be quiet and listen to the teacher. Aides don't stop overcrowding nor do they cure bad teaching (Leonard, 100, p. 45).

Or, try to imagine the consternation of the readers who encountered this claim for team teaching, "By tailoring lessons to the individual needs of the children, we have been able to help them achieve 100 per cent of their learning capacity, compared to 85 per cent or less in the traditional classroom" ("Team Teaching Proves Worthwhile in Chicago," 145, p. 284).

The causes of the confusion that has arisen concerning team teaching are not hard to find. A multiplicity of definitions of team teaching have appeared (Dean, 43), and many even operationally within the same school system (Evans and Womer, 49). The team organization has at times been confused with departmental and ability-group forms of school organization; several studies have been performed comparing the latter with the traditional self-contained classroom (Broadhead, 22; Heathers, 71; Hosely, 77; and Luchins and Luchins, 105). Multigrade approaches (Hull, 79; and Rehwoldt and Hamilton, 130) often have much in common with the team setting. No one definition of team teaching exists. The theories underlying team teaching (Brownell and Taylor, 25; and Fishler, 52) and the models of team teaching (Brownell and Taylor, 25; Lobb, Noall, and Slichenmyer, 102; and Singer, 133) that have been proposed illuminate many approaches to, and definitions of, team teaching.

The scope of projects involving teacher teams is extensive (Baynham, 15). In fact, team teaching has been tried on almost all levels. Three elementary team organizations that are well-reported exist at Englewood, Florida, (Gilchrist, et al, 58; and Goodlad, 61), Lexington, Massachusetts, (Anderson, 7; Anderson, 8; and Gilchrist, et al, 58), and Norwalk, Connecticut, (Anderson, 7; Becker, 16; Gilchrist, et al, 58; and Perkins, 125). At the secondary level, team projects are quite numerous (Singer, 134), especially in social studies and English. Very little has been attempted on the college level, although some studies exist here also (Hanvey and Tenenberg, 69; and Nelson, Hahn, and Robinson, 116).

Team teaching has even been experimented with in Sunday Schools (Foster, 56; and Keach, 89).

How has team teaching been received? The majority of the articles that have appeared take either an optimistic or pessimistic attitude toward team teaching; relatively few have attempted to present both sides of the issues involved. (However, see "Critical Look at Team Teaching, " 38.) Those articles favoring team teaching note many advantages: Flexibility in both teacher and pupil schedules is often noted (Empey, 48; Lambert, 97; Smith, 137; and Thomson, 146). Some articles claim superior achievement and personality gains for pupils in team organizations; others point out that gains in these areas would be more likely except for the fact that no teacher has yet been trained by a college or university specifically for team teaching (Cunningham, 39; and Fischler, 52).

Other alleged advantages may be listed:

- Team teaching recognizes the near-impossibility of a single teacher being well-prepared in a multitude of subject areas. Different teachers' specialties can compliment one another (Becker, 16; Dean, 42; and Varner, 152).
- Team teaching aids beginning teachers to "break-in" and helps experienced teachers to improve - an inservice function (Bloomenshine and Brown, 20; Elliott, 47; and Nagle, 115).
- Team Teaching draws from team members their best effort at independent thought and creativity (Andree, 12; Cooper, 35; Gross, 64; Heller and Belford, 72; Iannizzi and Robinson, 80; Varner, 152). (Also note a dissenting opinion,

Hanvey and Tenenberg, 69).

- Team teaching engenders a climate favorable for the development of a research attitude on the part of the teacher (Bloomenshine and Brown, 20; Haan and Adams, 66; and Johnson and Lobb, 84); a climate and attitude not ordinarily present (Singer, 134).
- Team teaching allows continuity of instruction, even when substitute teachers are used (Adams, 2; Marsh, 107; and Sweet and Dunn-Rankin, 141).
- Team teaching brings about a reduction of discipline problems.
- Team teaching allows the teachers to note better and to serve individual differences (Bloomenshine and Brown, 20; and McCurties and Schycker, 110); this concept is closely allied to the idea of vertical teams and non-graded approaches that rebel against "graded lockstep" (Brown, 24; and Gilchrist, 57).
- Team teaching allows and encourages independent study by the pupils and the establishment of student seminars (Glancy, 60; Harrison, 70; Heller and Belford, 72; Heymeyer and McGrew, 73; Lambert and Boyan, 98).

A common thread weaves through all of the articles by the proponents of team teaching, namely the potential created for effective staff utilization. Closely related to the concepts of staff utilization and team teaching, and in many ways inseparable from them, is the notion of employment of teacher aides. Education has often tried various plans for utilizing laymen, but never to such an extent as recently. Often given credit for reviving the recent teacher-aide plan is Greenshields (63). Aides are used to relieve teachers of nonprofessional duties (grading certain papers, collecting money, study hall duty, etc.); the aide is employed to alleviate the critically low level of the teaching talent reserve (Polos, 129). An extensive description of the teacher-aide plan has appeared (Park, 123) as well as a time and motion study that reports a markedly favorable shift in percentages of time the teacher spends on professional duties with implementation of an aide plan (Park, 122). Time freed for teacher preparation and other duties is reported at 20 and 25 per cent by other sources (Bleifeld, 18; and Noall and Jensen, 118). The number of available personnel in a community who wish to be aides is usually reported as surprisingly large.

But praise for team teaching and its allied areas is not unanimous. Certain writers have appeared to warn of the bandwagon nature of teacher teams (Bush, 28; Falk, 51; and Gilchrist, 57). Keliher (90) not only warns of the bandwagon effect but also expresses concern over the passivity of children that he believes team teaching will produce. Most other writers that are critical (Drummond, 45; Hoppock, 75; Ploghoft, 128; and Rzepka, 131) take issue with theoretical concepts of team teaching and have not been involved in any team teaching experiments. Only two sources (Cobett, 34; and Hanvey and Tenenberg, 69) report marked dissatisfaction with an experiment involving team teaching.

Substantial support for the self-contained classroom exists (see Hamalainen, 68; and Miel, 112). By "self-contained classroom," the modified variety is implied, not the pure variety in which only the teacher, and no specialist of any kind, teaches the pupils in a single classroom (Stiles, 139). Ackerlund (1) found about half the elementary teachers still prefer the self-contained organization over any other. But Anderson (9, p. 38), goes so far as to call the self-contained classroom "one of the sacred cows of education."

An interesting article (Fischler and Shoresman, 53) enumerates seven basic assumptions being challenged by the team teaching concept. Among these are the self-contained model and the idea that the ideal class size approximates thirty. The notion of the ideal teacher-pupil ratio is quite firmly entrenched in education; surveys conducted in 1945 by an educational magazine (156 and 157) indicated that over 90 per cent of elementary and secondary teachers favored classes under thirty pupils. More recently Kopp (94, p. 541) writes "A good individualized program in the elementary school requires a reasonable teaching load of approximately 20-25 children... In crowded situations, if necessary, let's curtail sessions and keep the groups small." The tendency to challenge the sanctity of the 1:25 teacher-pupil ratio can be noted (Andree, 12; and Bissex, 17).

Several points should be made here. First, team teaching does not necessarily increase or decrease the teacher-pupil ratio (Anderson, 7; and Smith, 137). Second, team teaching does not necessarily mean large group instruction, although Clark (32) points out that team teaching is becoming more and more identified with large group instruction; for example Wood (159, p. 30), writes "Usually team teaching isn't used with classes less than 100..." Some critics of the team organization, (Drummond, 45; and Hoppock, 75), seemingly base their remarks on the assumption that only large classes exist in the team.

Actually many schools follow the scheme of 40 per cent individual study, 20 per cent small group instruction, and 40 per cent large group instruction (Lord, 103). Anderson (8) reports that at Franklin Elementary School in Lexington, Massachusetts, only 20-25 per cent of school-day involves large group instruction while small group instruction is used for the remainder. A third point that should be made is that there is no consistent evidence supporting either the products of small or large class instruction in preference to the other (Ewan, 50; Goodlad, 62; Johnson and Lobb, 84; Johnson, Lobb, and Patterson, 85; Kidd, 92; Larmee and Ohm, 99; Nesbitt, 117; and Smith, 135), although a recent study involving high school seniors showed a class of 125 superior in all achievement areas over a class of 35 (Warburton, 153).

Another frictional area concerns the hierarchical arrangement of teachers on the team. Cashen (30) has wisely pointed out that team leadership should be given only to those with the credentials to warrant it. Friction arises over the wisdom of creating the hierarchy, which must be considered more than an administrative device (Tabachnick, 143). The opposing camps are readily identifiable on this issue. Favoring the hierarchy, with commensurate pay increases for the higher positions, are a host of educators (Anderson, Hagstrom, and Robinson, 10; Becker, 16; Bush, 27; Fischler, 52; Gross, 65; Lambert, 97; Lambert and Boyan, 98; Mayer and Woolridge, 108; Polos, 129; and Trump, 151). Opposing the hierarchy, Mahoney (103) calls for coordinate teachers of equal rank, while Drummond (45), Howe (78), and Weiss and Morris (155) are openly critical of the hierarchical model. King (93) advises no higher salaries for team leaders as he has sensed considerable teacher resistance to this (but certainly not from the leaders themselves). Another source (Tink, 150, p. 120) states, "If there is enough additional work and responsibility for the chairman of the teaching team, perhaps that member should be paid for the additional work, but not for so-called superiority of teaching." Ploghoft (128, p. 220), especially critical of higher salaries for better teachers, implies that the hierarchical model "borders on professional chicanery," and states, "First since the Lambert plan requires a team leader and teaching members who are less competent (that is why they are paid less) than the leader, the administrator must intentionally recruit some inadequate teachers or arbitrarily classify some as less competent than the leader." Perhaps Stoddard (140, pp. 156-157), has the most telling reply for those critical of the hierarchical arrangement:

> Let us face it. The classroom teacher is the last practitioner in the western world to respond to the industrial revolution. In the classroom there is practically no division of labor; apart from the textbook there is small utilization of equipment...
>
> Below the college level, all teachers are supposed to be equally replaceable parts... A teacher is supposed to make progress in the profession not through promotions in rank, with corresponding salary differentials, but simply by staying on the job... Monetary incentives are not the full answer to any personnel problem, but to reject them as irrelevant is to doubt the basis appeal of a capitalistic society.

Even among the proponents of team teaching, certain common problems are recognized to exist. Cost is one of these. Only two sources suggest that the team organization will cost less than the self-contained organization (Bruntz, 26; and Nesbitt, 117). Two others report no extra costs due to a team organization (Evans and Womer, 49; and Partridge, 124). Others report increased costs ranging up to 15 per cent, but usually this increase is due to the need for different types of materials rather than increased salary cost (Anderson 6 and 8; Becker, 16; Boutwell, 21; Lambert, 97; and Polos, 129). Another recurrent problem is that of selection of compatible team members. Caution and care are recommended by many

(Adams, 2; Clark, 31; Cunningham, 40; Fisher, 54; Heller and Belford, 72; Hemeyer and McGrew, 73; Iannizzi and Robinson, 80; King, 93; Loretan, 104; Marsh, 107; Mitchell, 113; Smith, 136; Stetson and Harrison, 138; and Stiles, 139). An additional problem, for the team members themselves, is the need for finding even more time for planning and evaluation meetings (Anderson 5; Iannizzi and Robinson, 80; Polos, 129; Sweet and Dunn-Rankin, 141; and Szabo, 142). McCallum, Henry and Nesbitt (109) bring up the problem of turnover of team personnel. *

Interestingly, a foreign observer (Dagnall, 41, p. 1080) views team teaching as advantageous to teachers for the following reason:

> It is said that the housewife with small children is one of the loneliest persons in our society. Could this also apply to some extent to our primary school teachers? If so, then team teaching may be of great value in preventing the feeling of teaching in isolation, and in replacing spasmodic contacts [with other adults] by considered discussion. Such discussion could lead to a raising of professional standards and a clarification of aims such as our education so urgently requires.

Well over a majority of the reports on team teaching indicate favorable attitudes toward it. Many authors approve the ideas of flexibility and staff utilization germane to the team teaching concept, among them Bloomenshine (19), Clawson, (33); Horn, (76); Howe (78); Keppel (91), and Mayer and Woolridge (108). One source ("Administrator's Guide to Team Teaching, " 3) points out strong administrative support for the method of organization, while team teaching, where tried, almost invariably is favorably received by a majority of the participating teachers (for example see Adams, 2; Crandell and Piel, 37; Gross, 64; Jenson, et al, 83; Jonsson, 87; Nesbitt, 117; or Noall and Rose, 120). Drummond (45, p. 163), however, brings up a good point, "What is essential for effective learning is not

---

*In some respects this turnover and need for constant orientation was especially noted in the case of this project's Intermediate Team.

necessarily highly correlated with what teachers prefer." Parent opinion about the team organization was sampled in three of the studies above (Crandell and Piel, 37; Gross, 64; and Jensen, et al, 83) and in each case was highly favorable. Student opinion, when sampled, was highly favorable (Adams, 2; Bloomenshine and Brown, 20; Gross, 64; Jensen, et al, 83; Johnson and Lobb, 84; Jonsson, 87; Noall and Rose, 120; Sweet and Dunn-Rankin, 141; and Varner, 152).

The wide support for team teaching is remarkable when one notes that the first team project was started as recently as 1957. Much impetus has been given to team projects by the School and University Program for Research and Development (SUPRAD), (See Hagstrom, 67), and the Ford Foundation (See Time, Talent, and Teachers, 149). Many school districts are so impressed by team teaching that they are building schools designed for this organizational approach (Anderson and Mitchell, 11).

Clearly, the phenomenon of team teaching is spreading, and some individuals already are predicting the disappearance of the self-contained organization (Adams 2; and Anderson, 9).

B. Interaction as Related to Team Teaching

Of great concern to most educators is the nature of the interaction that takes place within the classrooms of our schools. Until recently, interaction models capable of capturing this interaction process simply were not available. Recently, such models have appeared. In 1960, Lewis, Withall, and Newell (101) constructed a communication model for the study of classroom interaction; this instrument (hereafter referred to as the Withall Model) represents a modification of both

Bales' Interaction Process Analysis (13) and Withall's Index of Classroom Climate (158). Also appearing in 1960 was Flanders' Interaction Analysis (55). This model (hereafter referred to as the Flanders Model) is sensitive not only to the events, but uniquely also to the sequence of events taking place in the classroom.

A number of studies have been concerned with the interaction between teacher and pupil in various types of classroom organization (Anderson and Brewer, 4; Bank St. College of Education, 14; Brookover, 23; and Cornell, Lindvall, and Saupe, 36). A carefully controlled study comparing the interaction processes of the team and self-contained organizations, and making use of the communication models mentioned above, has not, as of yet, been reported in the literature.

However, a number of team projects have made subjective evaluations that are pertinent to the area of interaction. These evaluations cluster about two aspects of classroom interaction: student discussion and student discipline. Student participation must certainly be curtailed in the frequent large-size team classes, according to Hoppock (75); Keliher (90), as mentioned above, is concerned about the passivity of children under the team organization. Countering these, Heller and Belford (92, p. 109) point out that with a well-prepared master teacher the large class approach can be effective:

> This interaction between the teacher and the pupil is dependent upon the activity of the learner. To many educators this activity is often interpreted as physical activity which is focused upon the completion of some project. Such an interpretation, however, is shortsighted. The crucial factor in the learning situation for the learner is his mental activity. A pupil may be overtly very passive and docile, but his mind may be racing.

In actual experiments, Smith (136) reports more provision for discussion under team teaching while Dillion (44) reports less.

In the area of discipline, almost all sources report fewer discipline problems (the reader is reminded that these are subjective, not controlled and objective, evaluations); the sources include elementary, junior high, and senior high team projects (Bloomenshine, 19; Bruntz, 26; Giltinan, 59; Jonsson, 87; Loretan, 104; Marsh, 107; McCurties and Schycker, 110; Montague, 114; Noall and Jensen, 118; Sweet and Dunn-Rankin, 141; Taylor, 144; and Ward, 154). Only one source (Adams, 2) reported increased discipline problems using the team organization; these problems appeared on the second grade level but not on the fourth and sixth grade levels in the same project.

C. Personality and Achievement as Related to Team Teaching

Work of some early investigators (Pistor, 126; and Wrightstone, 160) suggest that whenever any thoughtful attempt is made to improve elementary education, gains are made in pupil attitude and adjustment, and there is usually equal accomplishment in skills. Securing adequate measures of the latter, achievement, is relatively much easier than securing objective measures of personality adjustment.

1. Personality

Ploghoft (128) expresses concern that the team organization waters down a strong and healthy identification of the pupil with a single self-contained teacher, an identification which is not possible in a team setting. In a more tangible vein, Crandell and Piel (37) report that elementary students taught by an experimental team were 9 per cent higher than self-contained students in "school-related problems" on a pupil problem inventory. Lambert (97), on the other hand, reports that most pupils benefit emotionally from having several teachers rather than one, although both he and Szabo (142) report that a few students must be separated due to confusion or marked insecurity in the new setting. (Szabo bases his comments on a high school team, Lambert on an elementary team.) On the other hand, other educators (Adams, 2; Jensen, et al, 83; Loretan, 104; and Smith, 135) report no ill-effects personality-wise to pupils taught under team organizations.

2. Achievement

Achievement is one of the primary areas of concern of projects comparing team and self-contained organizations, although Pitruzzello (127) has been critical of those who consider only differential achievement in their appraisals of team teaching. Holmes and Harvey (74), conclude, as a result of a study of arithmetic achievement, that an effective teacher, one sensitive to children's individuality, is a more influential factor than the type of grouping arrangement.

Although few carefully controlled studies have appeared in this area, most sources report no marked difference in achievement between team and self-contained organizations (Adams, 2; Anderson and Brewer, 4; Anderson, 6; Bloomenshine, 20; Crandell and Piel, 37; Dillion, 44; Johnson, Lobb, and Swenson, 86; Lambert, 96; Loretan, 104; Mahoney, 106; and Partridge, 124). Others report marked better achievement on the part of the team organization (Noall, Riggle, and Jensen, 119; and Noall and Rose, 120).

Only five of the sources above have direct reference to elementary school teams. Adams (2) reports especially favorable achievement for a fourth grade team, but Crandell and Piel (37) felt that the fourth graders in their school would benefit more from the self-contained organization. The other three sources (Anderson, 6; Lambert, 96; and Mahoney, 106) report no significant differences in achievement in the comparison of team and self-contained organizations. The literature does not, at this time, yield evidence that achievement is better under either organizational scheme, team or self-contained, on the elementary level.

IV. STATISTICAL METHODS

A. Analysis of Variance

The Analysis of Variance (ANOVA) has generally been used in the present study to determine what differences between the team and self-contained organizations are significant, that is, show differences in means between organizations greater than would be expected by chance.

The types of student or teacher behavior which are being analyzed by the Analysis of Variance method are normally called the factors of the analysis: The factors are the independent variables. Suppose, for example, one wishes to determine whether students do more talking in the team classroom or in the self-contained classroom, and whether there is a difference between the amount of talking by boys and girls, and whether this difference depends upon the difference between organizations. An ANOVA designed for this study would treat the type of organization (team or self-contained), and sex (boys or girls), as independent variables and the dependent variable would be the amount of talking.

Such an analysis would require four measurements on the dependent variable: the amount of talking girls do in the team, the amount girls do in self-contained classes, the amount of talking boys do in the team, and the amount boys do in the self-contained classes. This relationship between the independent and dependent variables is shown in the design table below. Such a design is called a 2x2 factorial division.

TABLE 1

Design of a 2x2 Factorial ANOVA of
Amount of Talking by Sex and Organization

Sex	Team	Self-Contained
Boys	$\overline{X}_1$	$\overline{X}_2$
Girls	$\overline{X}_3$	$\overline{X}_4$

There are four cells in this design which would show the means $(\overline{X}_i)$ for the amount of talking for each of the dependent variable situations.

An Analysis of Variance is computed on the observations in each cell to determine whether differences between the means of cells being compared is greater than could be expected by chance. The actual test is an F-Test. A corresponding table of critical F-values indicates the F-value required for significance at a specified level, say, less than five times in a hundred ($p < .05$). If the F-ratio equals or exceeds this critical value it is significant at the .05 level, and the probability that a significant difference between the means has occurred by chance is less than .05.

TABLE 2

F-Ratios for Amount of Talking
by Sex and Organization

Source	Amount of Talking
Sex	F_1
Organization	F_2
O x S	F_3

One notices that the ANOVA would yield three F-ratios for the supposed study. The first two of these F-ratios are for main effects, and show whether the independent variables indicate significant differences, of themselves, in the amount of talking. Thus if F_1 was significant, it would show that the amount of talking boys and girls do is significantly different, regardless of organization; and if F_2 was significant, similarly, it would show that the amount of talking between organizations is significantly different, regardless of sex.

The third F-ratio represents the interaction effect. An interaction effect shows whether the differences between the dependent variable is consistent with different effects of the independent variables. If F_3 was significant in our study; it would demonstrate that in one organization sex_A talked more than sex_B, but that in the other organization sex_A talked less than sex_B.

It will be noted, however, that in no cases do the significant main effects or interactions tell us the source of the significant effects.

In order to determine where the significances occur, once the F-ratio indicates they are present, one must go back to the cell means. In our example, again, the significant interaction effect did not tell us which sex talked more in which organization. We might go back to our cell means and find some such results as are graphed in Figure 1. According to this graph, boys do considerably less talking than girls in our imaginary team and considerably more in the imaginary self-contained classrooms. Such a cross-over in the graphs of the independent variables is generally reflected by a significant interaction effect, although normally the cross-over may be much less marked than in this imaginary case.

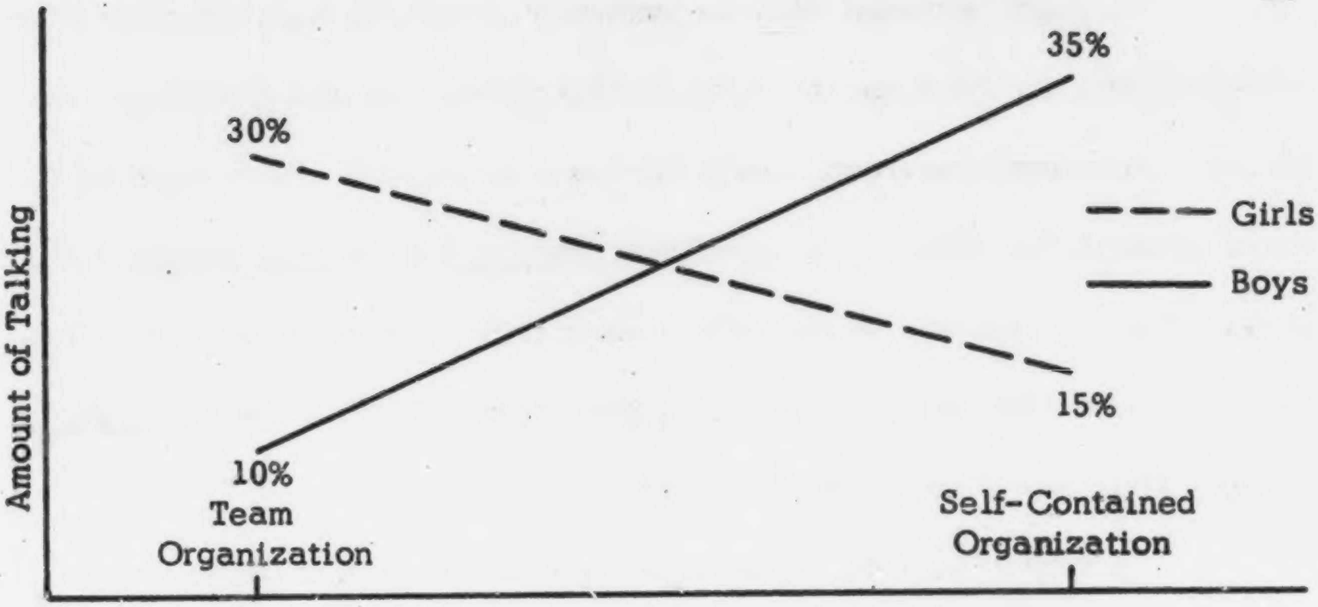

Figure 1: Amount of talking by sex and organization.

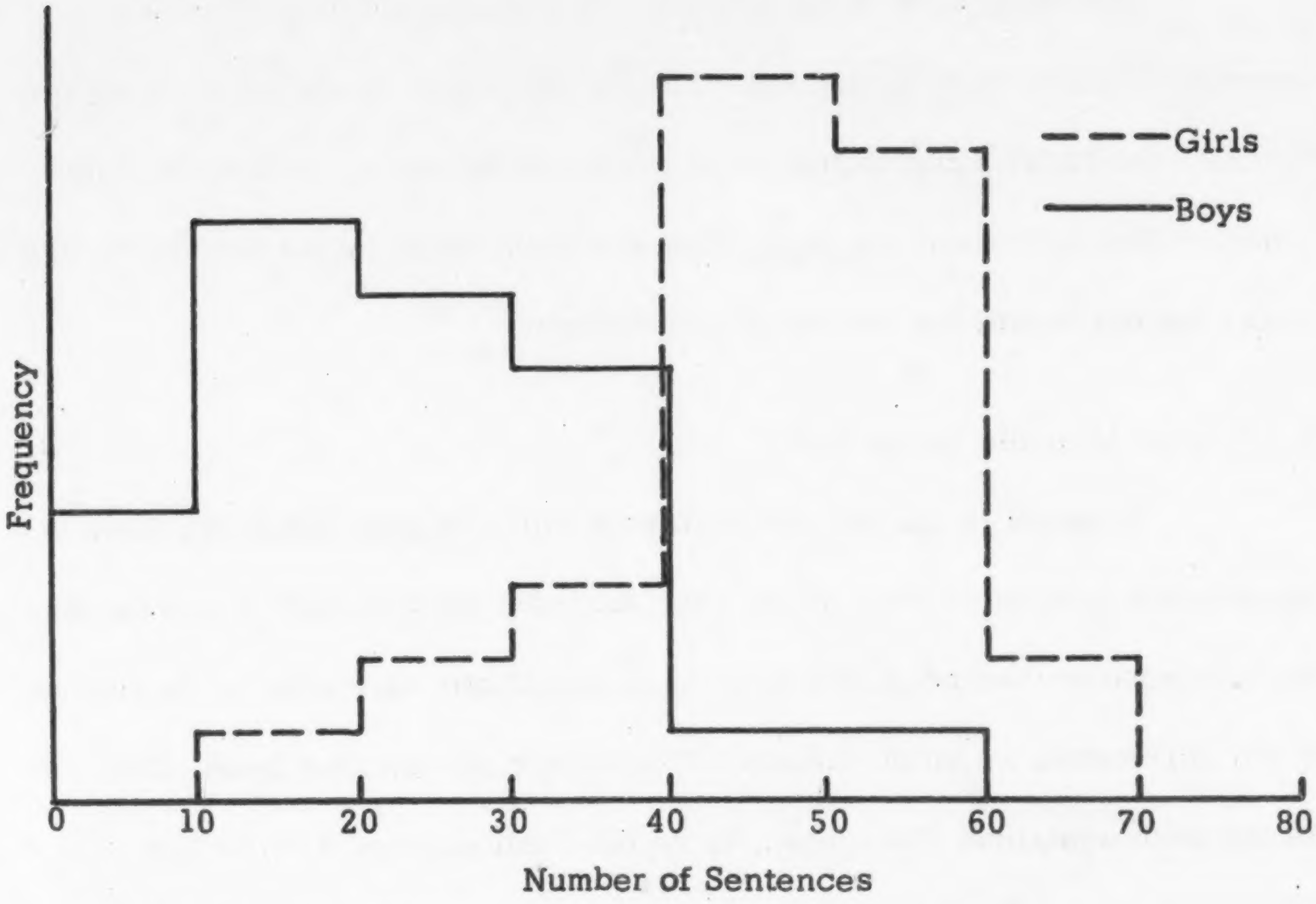

Figure 2: Frequency of number of sentences by sex.

It might be noted that the amount of talking in the imaginary team averages 20%, (if there are the same number of boys as girls), and 25% in the self-contained classroom, (with the same provision). Such a difference would probably be shown by a significant main effect for organization, (F_2) given sufficient numbers of students in each cell. But since girls and boys do the same amount of talking across organization (22.5%), the main effect for sex, (F_1), would not show significance.

Normally, as can be seen, one would not need to graph the cell means to show main effects. These would be immediately apparent from the results, and the source of significance could be seen by inspection of the cell means.

The designs of these studies are normally extensions of the factorial division to cover a greater number of cells. In one case, however, a special statistical computation is necessary to determine sources of significance. Whenever there are three organizations, mere inspection of the cell means cannot locate the source of significance.

B. Duncan Multiple Range Test

Suppose in our amount of talking study we had examined three organizations instead of two, let us say, adding a second team organization, and organization had been shown to be a significant main effect. Returning to the cell means we might discover 20% talking for the first team, 25% for the self-contained classroom, as before, and perhaps 21% for the additional team. The fact that organization is significant in this case, however, does not tell us whether the self-contained classroom is significantly higher than both team organizations, or only significantly higher than the first team, and it does not tell us whether the additional team is significantly higher than the first team.

In order to determine the precise source (or sources) of this significance, one may use the Duncan Multiple Range Test, a computation designed to show the minimum difference between two means which is required for the means to be significantly different. In the present paper, therefore, whenever an organizational main effect has been found to be significant, the Duncan range has been shown on the table of organizational means. A comparable table for our imaginary study is shown below.

TABLE 3

Mean Percentage of Talking by Organization

	Organization			Duncan Range
	Team 1	Team 2	Self-Contained	
Amt. of Talking	20	21	25	3

According to the illustrated Duncan range, there is significantly more talking in the imaginary self-contained class than either team organization, but there is not significantly more talking in Team 2 than in Team 1.

C. Bartlett's Test for Homogeneity of Variance

There are several assumptions which must be met for the analysis of variance. One is that the dependent variable has an underlying continuous, normal distribution. Secondly, independence among the dependent variable scores must be assumed. That is, in reference to our example, the amount of talk by student A is independent of the amount of talk going on by student B. Generally, these two assumptions can be met by logical design of the experiment.

The third requirement for an analysis of variance cannot be answered easily on logical grounds. This is that the divisions within an independent variable show homogeneity of variance, which means that they have approximately equal variance under the same conditions relative to the dependent variable. In the case of sex distribution in a random sample, it is assumed that variation in amount of talk of boys and amount of talk of girls is about equal.

In the case of classroom behavior, however, the criterion of homogeneity of variance is much more difficult to meet. Let us assume, for instance, that we wanted to study the amount of time teachers talk as well as the amount of student talk in our imaginary organizations. A single analysis of variance could be computed across all three variables, type of talk (student or teacher), organization and sex. This would be a 2x2x2 factorial division.

The question remains, however, whether it would be likely that there would be the same variation in teacher talk as in student talk, and this question would have to be answered before the 2x2x2 analysis could be undertaken. In order to find out the relationship of variances, a Chi-square test is computed across variances, according to the method known as the Bartlett's Test.

In the present project, precisely this question of the independence and homogeneity of variance of student and teacher behavior in the classroom was raised in the study of student-teacher interaction, and the Bartlett's Test was performed on the interaction data. According to these results, the types of student and teacher interaction did not show homogeneity of variance, and it was therefore necessary to use a separate analysis for each type of student-teacher interaction.

In our hypothetical case, this would be equivalent to finding that variation in student talk was not comparable to variation in teacher talk. If this had been demonstrated, then separate 2x2 factorial analyses would have to be performed on each type of talk, in place of the 2x2x2 analysis for both types ot talk.

D. Analysis of Covariance

Two of the analyses of the present project used the analysis of covariance method. This type of analysis may take into account more than one measure on the subject which may influence the dependent variable. In the imagined example, for instance, the only measure used in the analysis was a direct observation on amount of student talk. Suppose, now, we felt that student talk was highly related to student age to the nearest month, and we wanted to make the same study, taking into account the various ages of the students in the different organizations. It would certainly be impractical to introduce age as a new independent variable because there would undoubtedly be too many ages represented to allow a large enough sample in any one age group.

By the analysis of covariance technique it would be possible to adjust the measurements on student talk to take into account the variance due to student age. Once having made these adjustments, the measurements would represent, essentially, the amount of talk of the students, assuming they were all the same age to the nearest month.

The covariance analysis still yields F-ratios like the ANOVA. The difference is that it gives adjusted rather than raw means. Variables by which the means are adjusted are known as covariates: In the example, student age would be said to be "covaried out" of student talk.

E. The Kruskal-Wallis One-Way Analysis of Variance by Ranks*

Non-parametric techniques were required for analysis in some of the related studies. Non-parametric techniques do not require assumptions of normality. One of the most common is the chi-square test. Another used in this study is the Kruskal-Wallis One-Way Analysis of Variance by Ranks.

The chi-square test tests the hypothesis that two sample distributions come from the same population. In contrast to the parametric methods, the chi-square test tests the difference in distributions rather than the difference in means.

In keeping with our example, suppose we took our measures of the amount of talk for boys and girls and broke them into two distributions - one for boys and one for girls. An arbitrary division of say 10 sentences could be used for distribution intervals. Suppose these hypothetical distributions were as presented in figure 2 above. A chi-square test would then test the probability that these two distributions come from the same population. A significant chi-square ($p < .05$) would indicate that the probability that the sample distributions come from the same population is less than .05.

---

*This more complicated method of analysis was used in only one experiment; a side study of student behavior.

The Kruskal-Wallis One-Way Analysis of Variance by Ranks tests the probability that two or more independent samples come from different populations. This technique tests the null hypothesis that the samples come from the same or identical populations with respect to averages. It requires simply ordinal measurement of a variable that has an underlying continuous distribution.

The Kruskal-Wallis Test requires a ranking of the N scores from 1 to N. These are then combined in a computation to give the "H" statistic (Siegel, 132, pp. 184-193).

V. POPULATIONS AND EXPERIMENTAL DESIGN

Three different classroom organizations were studied over the two-year period. Two of these were the team and self-contained organizations, both in Washington School, and the third was the self-contained organization at Longfellow School. There were approximately 350 elementary school pupils in the three organizations, about 60% of the pupils, in grades 1-6 at Washington School, and the remaining 40%, in Longfellow School. Naturally the number of pupils changed slightly from year to year.

A. Washington School: Two Organizations

The Washington Elementary School is located at 545 West Dayton Street, in Madison, Wisconsin, in an economically depressed area. The majority of the parent population is made up of lower-middle class working people. Twenty percent of the families are currently receiving aid under the Public Assistance Program. At the other end of the spectrum, about ten percent of the parents are in managerial and professional occupations.

Before the initiation of this experiment, there were two classes in each grade of the school with about 20 pupils per class. The experimental design separated the pupil population at Washington, grade by grade, into two sections, an experimental organization and a control organization. Children were assigned to one of these organizations by the use of a table of random numbers.

1) The Washington School Self-Contained Organization

The control organization continued a modified self-contained classroom arrangement. The modified self-contained classroom was similar to the traditional self-contained classroom with the exception that the children were also exposed to specialists in the fields of physical education, art, and music. In this section of the school there was one class at each grade level, first through sixth.

2) The Washington School Team Organizations

In the other section of Washington School were two teams, each composed of a team leader, a regular teacher, two graduate interns, and a half-time instructional secretary. These teams corresponded to grades 1-3 and 4-6, and are referred to as the Primary and Intermediate Teams, respectively.

In the team organization all groupings were made according to the demands of each subject, and each child's individual potentialities and interests were considered. In social studies and science, for example, there was a great deal of large-group instruction, with student panels, debates, movies, and visiting resource speakers. In addition there were smaller follow-up, special interest, discussion and work groups which varied in composition according to the material which had been presented. A major concern of the teams was the development and reorganization of a social studies curriculum.

The faculty members were given a choice of whether they wanted to remain in Washington School and to help try out new methods of teaching, or whether they preferred to be transferred to another Madison public school. Only six out of twenty-six staff members chose to be transferred.

The interns were typically people who had gained their first degree in some field other than education and were working towards a master's degree in elementary education. They usually had no school teaching experience prior to their semester internship except for the previous summer session during which they participated in the University Laboratory School.

Each semester two new interns were integrated into the on-going instructional program as quickly as possible. Their time and activities were divided equally between the two experienced teachers.* This meant that in a number of subjects they changed places half-way through the semester. For example, one of the interns might have begun by working with the teacher responsible for reading. He would stay the first eight weeks. At the end of the eight-week period, the intern would join the teacher responsible for arithmetic. The interns participated in all teaching duties and responsibilities.

The instructional cost was about the same for the team and self-contained organizations, since in each team one regular teacher's salary was eliminated. This money was divided among the interns, instructional secretary, and team chairmen: Each intern received approximately $1,000 per semester; the half-time instructional secretary for each team received $1,500 per year; and the team chairman received approximately $1,000 per year above his regular salary in recognition of the added responsibilities of his position.

Certain difficulties arose in preserving stability in the Intermediate Team during the two years of the experiment. The original master teacher dropped from the team in the course of the first year due to pregnancy. As a result, communication problems arose between the remaining teachers. In the second year, a new team chairman and master teacher came on the team, who entered while their notion of the team concept was still being developed.

---

* Experienced teachers did not teach all subjects.

B. Longfellow School

The second control organization was incorporated in the experiment to measure possible contamination of variables in Washington School. Furthermore, measurements on the second school gave greater power to the statistical analysis. Longfellow School is nearby and comparable in terms of intelligence, achievement, socio-economic background, administration, and size of pupil population. The Longfellow Elementary School is located at 210 South Brooks Street in an area which is undergoing considerable urban redevelopment. The majority of the pupils' parents are from the lower middle class. Approximately 70% of the parents have jobs which would be considered unskilled or semi-skilled tasks. Approximately 25% of the families are currently receiving aid under the Public Assistance Program. About 10% of the parents come from the middle class and have skilled jobs.

In Longfellow School there were six grades with about 30 students in each class.

C. Teacher Bias in the Study

Care was taken to introduce bias in the selection of staff for this study: Only strong advocates of the team and self-contained organization were retained or employed.

Each group was stimulated by university consultants and the introduction of new teaching aids, that is, overhead projectors, teaching machines, and tape-recorders. All groups were stimulated to counteract possible Hawthorne effects.

D. Intelligence Quotient in the Organizations

I.Q. tests were administered before the experiment to make certain that the three organizations were evenly matched. The average I.Q.'s of the students is shown below in Table 4.

TABLE 4

Average I.Q.'s by Sex, Grade, and Organization*

		Organization		
Sex	Grade	Washington Team	Washington Self-Contained	Longfellow Self-Contained
Girls	2	107.3	103.8	109.7
	3	110.4	108.0	112.4
	4	109.2	105.2	115.1
	5	104.1	104.4	101.7
	6	102.5	97.4	104.9
Boys	2	101.9	100.8	101.0
	3	116.4	112.0	109.4
	4	113.8	112.4	110.4
	5	107.2	101.2	108.8
	6	105.0	105.0	104.2
Combined	2	104.4	102.1	106.3
	3	113.8	110.2	110.4
	4	111.3	108.8	112.2
	5	105.7	103.4	105.5
	6	103.8	100.6	104.6
All Grades		107.8	105.4	107.5

* _California Short Form Test of Mental Maturity,_ 1957.

E. Summary of the Project

Each of the general and related hypotheses stated as objectives in Section II above (page 5) was tested in an independent experiment, and several of these experiments were broken down into specialized studies. The total Team Teaching Project included, therefore, three main studies, Interaction, Personality, and Achievement, and four side studies, Teacher Awareness, Peer Stability, Absenteeism, and Student Behavior. The three main studies include a number of sub-studies, using different interaction models, different types of tests, and different experimental treatment groups.

Due to the complexity of the total project, this present report has been divided into three further parts, in each of which, data and methods of analysis are individually treated: Part Two covers the entire Interaction Study; Part Three includes the Personality and Achievement Studies treated separately; and Part Four includes the four side studies which, again, are treated separately. Finally, a brief discussion of the entire Project has been presented in Part Five of the report, with a recapitulation of the conclusions of the independent studies.

PART TWO

CLASSROOM INTERACTION

PRECEDING PAGES BLANK-NOT FILMED

I. INTRODUCTION

A. Classroom Interaction Models

Two models of classroom interaction, the Flanders and Withall Models, (see pp. 15-17), were used with certain modifications in the Interaction studies. The present section is intended as an introduction to these models, while the modifications of observation and categorization used in the study are presented separately in Sections II and III, following a summary of the study.

1) The Withall Model

There are fourteen mutually exclusive categories of classroom verbal behavior in the Withall Model, including all types of verbal behavior in which the teacher participates.

A detailed description of the categories of the Withall Model, with examples, may be found in Appendix A, but these may be summarized as follows:

1. Asks for Information
2. Seeks or Accepts Direction
3. Asks for Opinion or Analysis
4. Listens
5. Gives Information
6. Gives Suggestion
7. Gives Directions
8. Gives Opinion

9. Gives Analysis
10. Shows Positive Feeling
11. Inhibits Communication
12. Shows Negative Feeling
13. No Communication
14. Perfunctory Agreement or Disagreement

The Withall Model was designed to be used either by an observer in a classroom or with tape recordings of classroom interaction. The observer is to make the judgments about each unit of behavior, as it occurs, in terms of the dominant intent of the initiator. Both verbal and non-verbal aspects of behavior, where detectable, are used in making the judgment. The interval of ten seconds is recommended as the time unit. Observations are made sequentially and at least one category is assigned to each ten-second interval. When there is a shift in the inferred intent of the communicator within any ten-second interval, the interval is assigned additional category scores.

Upon completion of the recording period, the observer tabulates the frequency of occurrence of each category.

2) The Flanders Model

The Flanders Model is made up of ten categories, but these are not restricted as in the Withall Model, to the type of verbal behavior which includes only the teacher. The Flanders categories are also mutually exclusive and cover all types of behavior in the classroom. A detailed description of these categories is supplied in Appendix A, and they may be summarized as follows:

1.	Accepts Feeling	Teacher Talk
2.	Praises or Encourages	
3.	Accepts or Uses Ideas of Student	
4.	Asks Questions	
5.	Lectures	
6.	Gives Directions	
7.	Criticizes or Justifies Authority	
8.	Response	Student Talk
9.	Initiation	
10.	Silence or Confusion	

Trained observers categorize the classroom interaction every three seconds. This yields a series of observations which are tabulated on a 10x10 matrix whose rows and columns correspond to the categories. Observations are tabulated on this matrix by pairs, the row index indicating the first member of a pair, and the column index indicating the second member of a pair. To make a series of observations complete, it must begin and end with the same number. This has been arbitrarily selected as the number 10. Thus, the series of observations, 10-4-8-2-10 might indicate a twelve-second period in which the teacher had asked a question, a student had answered, and the teacher had praised the student. This interaction would be tabulated on the matrix as shown in Figure 3.

	1	2	3	4	5	6	7	8	9	10
1										
2										X
3										
4								X		
5										
6										
7										
8		X								
9										
10				X						

Figure 3: Flanders matrix, showing tabulation of a typical twelve-second classroom interaction.

It will be noted that summation of either the rows or columns reveals the frequency of responses within each category, (disregarding the additional arbitrary 10). Thus data can be recorded in two inter-related ways: By actual frequency for each category, which leads to the statistical analysis, and by matrix tabulation, which leads to a descriptive analysis.

B. Schedule and Outline of the Interaction Project

Three independent studies of student-teacher interaction in the team and self-contained classrooms of Washington School and the self-contained classrooms of Longfellow School were conducted during the first and second semesters of the academic year 1961-62, and the first semester of the academic year 1962-63. These studies used the Withall and Flanders Models of classroom interaction*, or variations of these models, and two observational procedures, as shown in Schedule 1, below.

SCHEDULE 1

Interaction Studies, 1961-63

Study	Observation Period	Observation Procedure	Interaction Model
1st Semester 1961-62	Week of January 8	Remote station tape recorder	Revised Withall and Revised Flanders
2nd Semester 1961-62	Week of June 4	Remote station tape recorder	Revised Withall alone
1st Semester 1962-63	December 4-January 24	Within class observers	Flanders alone

* See Page 15 above.

INTERACTION PROJECT, 1961-63

1) First semester 1961-62

At the outset of the project, emphasis was put on an estimation of the comparative analytic powers and differing areas of sensitivity of the Flanders and Withall interaction models. In this study two particular questions were raised: 1) How sensitive are the models to basic grade level differences in teaching methods? 2) Are categories of student-teacher interaction within the interaction models independent?

The first question was raised because student-teacher interaction might vary more greatly between the primary and intermediate grade level than between the team and self-contained type classroom organization. The main effects of grade level and organization, and their interactions, were estimated by an Analysis of Variance.*

The question remained, whether in conducting an Analysis of Variance the categories of interaction should be treated separately or could be included in a single Analysis of Variance across categories. The homogeneity of variance of the interaction category frequencies was tested by the Bartlett** method.

Furthermore, different lengths of observation intervals were used on the two interaction models.

---

*A "layman's" description of this statistical technique is given in Part One, IV, pp. 20-24.

**The Bartlett Test for Homogeneity of Variance is described in Part One, IV, pp. 25-27.

For the Withall Model, observations were taken at intervals of twenty-seconds, which was considered a long enough interval that an observation would be likely to be independent from the previous and following observations, while observations on the Flanders Model were taken every three seconds, as Flanders originally recommends. To compare the effects of using different models, the same class sessions were observed with both models.

An observation period of thirty minutes was chosen since both interaction models were used on the same class sessions. This period seemed a reasonable compromise between the entire class session, normally used with the Withall Model, and the twenty-minute period recommended by Flanders.

Finally, categories were simplified on both interaction models. It was felt that a grouping of closely similar categories would facilitate the examination of category independence. Thus, for instance, Categories 1-3 of the Flanders Model which cover the teacher's acceptance of student feeling, praising or encouraging, and acceptance of student ideas, were reduced to one category. The revised models clearly separated types of classroom interaction into distinct categories.

2) Second Semester, 1961-62

Analysis of the results of the Revised Withall Model observations, which had been spaced at twenty-second intervals, demonstrated that classroom interaction categories did not show homogeneity of variance and therefore that the analysis of variance must be conducted on individual categories, even when observations were spaced at distant intervals.

The first semester study, however, pointed to a relatively insignificant effect of grade level in the distribution of student teacher interaction categories.* Out of a total of fourteen categories for both models, three categories showed grade level significance, while six categories were significant for organization. Nevertheless the Revised Withall Model, in showing two out of six categories significant for grade level, seemed more sensitive to this variable than the Flanders Model, which detected only one out of eight categories significant for grade level.

In the second study a closer look was taken at grade level significance. The Revised Withall Model was used again, but grade level was broken down into individual grades, rather than into the basic levels corresponding to teams, primary and intermediate, which had been treated in the first study.

Furthermore, the effect of subject was investigated, and in order to isolate this effect from possible organizational contamination, only self-contained classrooms were observed. Total class sessions on a particular subject, sometimes as long as ninety minutes each, were observed rather than only thirty-minute sessions, as in the previous study.

---

* However, the descriptive analysis of classroom interaction of the Flanders Matrix tabulations in the second year indicated that grade level differences may have a greater effect between intermediate and primary teams than between the intermediate and primary levels of the self-contained organizations.

3) First Semester, 1962-63

While subject area was shown to be significant for nearly all categories in the second study, grade level proved to be nonsignificant in all categories. In the final study, therefore, in which all three organizations were observed, subject matter was again treated as an independent variable.

For this study, however, the Flanders Model was used, in order that the results of the Withall Model from the previous studies might be replicated on a different model. However, the model was used with the normal categories in order to study the sequence of interaction patterns within the classrooms. Finally, in this study, periods of observation were reduced to twenty minutes in accordance with the normal Flanders procedure.

A descriptive analysis of classroom interactions according to the Flanders Matrix process was undertaken in addition to the statistical analysis in the final study to observe more closely organizational and grade level effects which might not show statistical significance.

II. OBSERVATION

A. Method of Observation

Classroom observations are either made directly in class, or from tape recordings.

In the present project, both methods of observation were used. During the first year of study, class sessions were tape-recorded: Teachers were given radio frequency remote control microphones, which they wore around their necks, and sessions were tape-recorded at stations remote from the classroom.

Tape recordings were used on the first semester study because both interaction models were to be applied to the same class sessions, and it was felt that a tape recording would make the initial conditions for categorization with the different models the same. Although only the Withall Model was used in the second study, tape recordings were again used, in order that an additional observational effect might not be introduced which would distort comparison of first and second semester results.

In the third semester of study, when the Flanders Model was used, observations were made by trained observers who remained in the classroom during the sessions, and categorized the student-teacher interaction as it was taking place.

B. Schedule of Observations

1) First Semester, 1961-62: Week of January 8, 1962

During this week, observation periods were randomly selected from the total teaching time of 20 participating teachers, with the following restrictions: (1) No two periods could be concurrent within any school building since the radio frequency remote control microphones for recording would interfere with each other at short distances; and (2) Each of the three organizations was observed an equal number of half-hour periods. The distribution of observation periods per teacher is shown in Table 5.

TABLE 5

Distribution of 30-Minute Interaction
Observations Per Teacher

First Semester, 1961-62

Teams	Washington Team	Washington Self-Contained	Longfellow Self-Contained	Grades
Primary	3 3 3 3	4 4 4	4 4 4	1 2 3
Intermediate	3 3 3 3	4 4 4	4 4 4	4 5 6
Observations per Organization	24	24	24	
TOTAL N		72		

Twenty-four half-hour observations were conducted in each organization. Since there were four teachers in each team, each team teacher was observed three half-hour periods. Each of the 12 teachers of self-contained classrooms, one for each grade level in each school, was observed four half-hour periods. In all, there were 72 observations.

2) Second Semester, 1961-62: Week of June 4, 1962

In this week, 12 teachers of self-contained classrooms, one for each grade level in each school, were observed. However, no team teachers were observed, and rather than limiting observation periods to a half-hour, complete sessions were recorded, ranging from 30 to 90 minutes each. Furthermore, observations were balanced for four subject areas within organizations: Six periods each were blocked for arithmetic, reading, science, and social studies. Restrictions pertinent to the tape recording procedures were followed, as in the previous observation period. Distribution of observations in this study is shown in Table 6.

TABLE 6

Distribution Of Interaction Observations
Per Teacher

Second Semester, 1961-62

Grades	Washington Self-Contained	Longfellow Self-Contained
1	4	4
2	4	4
3	4	4
4	4	4
5	4	4
6	4	4
Observations per Organization	24	24
TOTAL N	48	

3) December 4 - January 24, 1962-63

During the final observation period, observations were blocked for three, rather than four, subject areas. (The limitation to three subject areas in this study rather than the four subjects of the previous study was necessary because of complications in observation schedules for within class observations.) Eight periods each, of reading, arithmetic, and social studies, were observed for each organization. As in the observation period of January 8, 1962, 20 teachers were observed, and observations per teacher followed the same breakdown, shown above in Table 5. Thus there were 72 observation periods in this study.

As was pointed out in Section A, above, these observations were conducted within the classrooms. Trained observers categorized interaction according to the Flanders Model every three seconds for a period of 20 minutes. The 400 observations were entered by category number on a data matrix.

C. Observer Training and Observer Reliability

Training of observers and computation of observer reliability depended on two factors: the method of observation, and, the interaction model. Thus it is important to remember that observations which were made upon tapes that had been previously recorded in the classroom must be distinguished from observations which were made directly in the classroom. In the first year of the study, observers were trained to categorize student-teacher interaction from tape recordings. Final interaction data was not obtained until the observers had listened to the tapes. In the study of December 4 - January 24, 1962-63, observers were trained to categorize interaction in the classroom, and final interaction data was obtained directly from classroom sessions.

Schedules of observer training and methods of computing observer reliability may now be described relative to the methods of observation:

1) First Semester, 1961-62

Four graduate students were trained as observers in this study. Under the direction of a trained observer from the Withall Project,* two students assigned to the Withall Model learned the definitions of the 14 categories, and practiced in excess of fifteen hours on tape recorded class sessions, and spent one day observing classroom interaction in a school not participating in the study. Throughout the training period a constant check of inter-observer reliability was made by having both trainees make observations simultaneously and checking their percentage agreement (total number of observations in agreement per category summed over all categories and divided by total number of observations). At the end of the training period observers averaged 97% agreement. Inter-observer reliability was maintained throughout the period of observation by regular reliability checks.

For training, the two observers using the Flanders Model listened to tapes of ten class periods of 30 minutes each. Reliability was calculated in the same manner that was used with the Withall observations, and ranged from 81% to 93% with a mean of 88% for the ten training sessions. The difference in reliability level between Flanders and Withall observations is due to the higher frequency of Flanders observations. (Observations were taken every 3 seconds for the Flanders Model and every 20 seconds for the Withall Model.)

---

* N.I.M.H. Project 2M-6624, (111).

2) Second Semester, 1961-62

The same two graduate students who worked with the Withall Model in the first semester served as observers the second semester. In addition to these two observers, two trained observers from the Withall Project also listened to the tapes. The four observers were assigned at random to the recordings of complete class sessions of a subject area.

Although no training sessions were considered necessary, observer reliability was periodically checked for percentage agreement in the manner described above, and it ranged from 93% to 97% with a mean of 95% during the period of data collection.

3) First Semester, 1962-63

Two graduate students who had not participated in the previous studies were used as observers. These observers trained during the period, November 19-30, 1962. The training consisted of listening to five tapes and observing five classrooms jointly.

The reliability between observers was computed using the "Scott's Coefficient" as suggested by Flanders (55). This coefficient is referred to as "pi" and is computed by the formula:

$$\pi = \frac{Po - Pe}{1 - Pe}$$

where Po = proportion of agreement;

Pe = proportion of agreement expected by chance; and,

$$Pe = \sum_{i=1}^{k} p^2$$, where k = number of categories, and

p = proportion of tallies falling into each category.

The reliabilities during the training period ranged from .60 to .88 with a mean of .75. The higher reliabilities came toward the end of the training period. Three reliability checks were made at intervals during the data gathering period. These reliabilities were: .82, .80, and .91.

The drop in reliability between the Flanders observations in the first study and those in the third is probably caused by the fact that the former were taken from tape recordings, and the latter were made in the classroom. In the classroom, observers located in different parts of the room are more likely to have differing interpretations of interaction than observers who are listening to the same tape, and may replay segments in question.

III. DATA

Revised forms of the Withall and Flanders Models were used to categorize the interaction which had been tape recorded in the first and second semesters of the first year.

1) First Semester, 1961-62

The Withall categories (see pages 39-40) were reduced to six, as follows:

1: (1, 2, 3):	Asks for Information; Seeks or Accepts Direction; Asks for Opinion or Analysis
2: (4):	Listens
3: (5, 8, 9):	Gives Information; Gives Opinion; Gives Analysis
4: (6, 7):	Gives Suggestion; Gives Directions
5: (13):	No Communication
6: (10, 11, 12, 14):	Shows Positive Feeling; Inhibits Communication; Shows Negative Feeling; Perfunctory Agreement or Disagreement

The trained observers listened to the tapes and categorized interaction every twenty seconds. This yielded 90 observations per thirty-minute session; 72 of these observations were picked at random from each session, and frequency within category was computed for final analysis.

A revised form of the Flanders Model was also used, in which Categories 1-3 were condensed into one category, and Categories 5 and 7 were redefined in order to assure clearly independent meanings. These revisions were based upon analogy of the Flanders Categories to the Revised Withall Categories.

It will be noticed that the Withall Model distinguishes between strong and mild criticism while the Flanders Model does not. Withall Categories 10, 11, 12, and 14, combined in Revised Category 6, cover strong criticism, while Withall Category 8 (Gives Opinion) may cover mild criticism. In fact, it is quite difficult to distinguish giving information from giving opinion and analysis in the normal classroom lecture which often contains all three. Thus Withall Category 8 is most clearly applicable to a teacher's critical opinion of a student's response.

By analogy to Revised Withall Category 6, Flanders Category 7 (Criticizes or Justifies Authority) was used only in cases of strong criticism. Criticism of an unsevere nature, on the contrary, was classified with Flanders Category 5 (Lectures), by analogy to Revised Withall Category 3. Thus all forms of direct teacher communication except explicit direction and strong criticism were covered by Revised Flanders Category 3. The revised Flanders categories were, then:

1. Accepts Feeling, Praises or Encourages, Accepts or Uses Ideas of Students
2. Asks Questions
3. Lectures, Criticizes or Justifies Authority
4. Gives Directions
5. Strong Criticism
6. Student Response
7. Student Initiation
8. Silence or Confusion

The same tapes were categorized according to the Revised Flanders Model, independently from the Withall categorization. However, observations were made every three seconds on the Flanders Model while they were made every twenty seconds with the Withall Model.

The number of observations per category were changed to percentage scores for final analysis.

2) Second Semester, 1961-62

The same Revised Withall Categories were used upon the second semester tapes: Observations were made every 20 seconds, and 72 observations were picked at random from each session. Frequency within category was computed for final analysis.

3) First Semester, 1962-63

The observations of this study, which had followed the normal Flanders Model, and had been made in class, were transferred directly from the data sheet to the Flanders Matrix, according to the procedure described on page 41 above. Frequency per category was also computed from the Flanders matrices for final analysis.

The interaction data obtained from the three studies may be summarized:

1) First Semester, 1961-62:

   a) Revised Withall: Frequency per category
   b) Revised Flanders: Percentage per category

2) Second Semester, 1961-62:

   Revised Withall: Frequency per category.

3) First Semester, 1962-63:

- a) Flanders: Matrix Tabulation.
- b) Flanders: Frequency per category.

IV. STATISTICAL ANALYSIS

A. Statistical Methods

Homogeneity of variance of the frequencies of categories was computed according to Bartlett's Test of Homogeneity of Variance.

The significance of categories of observed interaction was analyzed by the Analysis of Variance (ANOVA). By analyzing each category separately the underlying assumptions for the ANOVA were met. A factorial ANOVA was used, having organization as one independent variable, in combination with one other variable.

When a significant organization effect was found, a Duncan Multiple Range Test was used to locate the sources of significance.* The Duncan Range was computed by the procedure outlined by Edwards (46). This procedure for testing the three organization frequencies would yield three between organization means, one for each combination of two means. However, since the organization frequencies were equal and the minimum gaps required for significance were very close, only the largest of the three minimum gaps is reported in the tables.**

B. Results

1) First Semester, 1961-62

The results of the Bartlett Test for Homogeneity of Variance, which was computed on the Revised Withall Categories, yielded a corrected Chi-square value of 82.33 which was significant at the .001 level (df = 4). Thus the categories did not show homogeneity of variance, and the analyses of variance were computed for each category rather than across categories.

---

*See pages 24-25.

** This yields the most conservative statistical estimate.

a) Revised Withall: Frequency per Category

The mean number of frequencies of the Revised Withall Categories by organization and grade level is shown below in Table 7.

TABLE 7

Mean Number of Frequencies of Revised Withall Categories by Organization and Grade Level

First Semester, 1961-62

	Washington Team		Washington Self-Contained		Longfellow Self-Contained	
Category	Prim. *	Inter. **	Prim.	Inter.	Prim.	Inter.
1	10.1	9.2	9.8	11.6	15.3	14.9
2	18.9	13.3	23.9	20.4	16.0	16.5
3	19.8	28.6	10.6	18.4	17.9	20.6
4	6.7	6.3	7.7	4.6	8.3	5.7
5	13.5	10.3	16.2	14.8	11.7	11.8
6	3.0	4.3	3.8	2.2	2.8	2.5

* Prim. = Primary ** Inter. = Intermediate

A 2x3 factorial ANOVA was computed for each category of the Revised Withall Model. First through third grade, and fourth through sixth grade, of the self-contained classrooms were treated as two grade levels respectively corresponding to the Primary and Intermediate Teams. The breakdown of this ANOVA is shown in Table 8.

TABLE 8

Division of 2x3 Analysis of Variance for Revised Withall Categories

First Semester 1961-62

Grade Level	Washington Team	Washington Self-Contained	Longfellow Self-Contained
Primary Level Grades 1-3			
Intermediate Level Grades 4-6			

Revised Category 6 (Shows Positive Feeling, Inhibits Communication, Shows Negative Feeling, Perfunctory Agreement or Disagreement) was dropped from this analysis since the frequency in the category was relatively very small, as can be seen in Table 7.

The resulting F-ratios of the ANOVAs for the remaining five revised categories are given in Table 9.

TABLE 9

F-Ratios of Revised Withall Categories by Category and Source

First Semester, 1961-62

	Category				
Source	1	2	3	4	5
Organization (0)	6.1**	3.63*	6.21**	-	2.87
Grade Level (L)	-	1.89	8.28**	5.88*	1.07
O x L	-	-	-	-	-

*p<.05
**p<.01

Note: In this and all succeeding F-value tables, a - indicates an F<1.

In Revised Categories 1, 2, and 3, the effect of organization was clearly significant. The sources of these significances can be seen in Table 10 which shows the breakdown of mean frequencies by organization. The Duncan minimum range has been computed in order to determine where the significance lies between organizations.

For Revised Category 1, (Asks for Information, Seeks or Accepts Direction, Asks for Opinion or Analysis), there was not a significant difference between the Team and Washington School Self-Contained classes; however, Longfellow classes were significantly higher in this category than the classes of the other two organizations.

TABLE 10

Mean Number of Frequencies of Revised Withall Categories By Organization

First Semester, 1961-62

Category	Washington Team	Washington Self-Contained	Longfellow Self-Contained	Duncan Range
1	9.7	10.7	15.1	3.4
2	16.1	22.2	16.3	5.4
3	24.2	14.5	19.3	5.8
4	6.5	6.1	7.0	
5	11.9	15.5	11.6	

Washington Self-Contained classes were significantly higher than both other organizations in revised Category 2 (Listens).

In Revised Category 3, (Gives Information, Gives Opinion, Gives Analysis) the Team was significantly higher than the Washington Self-Contained classes, while there was no significant difference between Washington and Longfellow Self-Contained classes.

The significance of grade level in Category 3 was caused by the fact that the intermediate levels were much higher than the primary levels; in revised Category 4 (Gives Suggestions, Gives Directions) on the contrary, the primary levels were higher than the intermediate levels. This reversal of frequencies can be seen in the means shown in Table 11.

TABLE 11

Mean Number of Frequencies of Revised Withall Categories by Grade Level

First Semester, 1961-62

Category	Grade Level Primary	Intermediate
1	11.8	11.9
2	19.6	16.7
3	16.1	22.5
4	7.6	5.5
5	13.8	12.3

b) Revised Flanders: Percentage per Category

Table 12, below, shows the mean percentage scores of the Revised Flanders Categories by grade level and organization.

TABLE 12

Mean Percentage Scores on Revised Flanders Categories by Grade Level and Organization

First Semester, 1961-62

	Primary			Intermediate		
Category	WT*	WSC**	LSC***	WT	WSC	LSC
1	7.7	9.1	9.5	9.8	6.0	8.0
2	13.5	14.5	18.2	11.6	16.2	18.9
3	18.7	7.2	12.7	23.2	14.2	17.2
4	14.5	12.5	14.8	15.5	15.2	13.3
5	1.7	2.3	4.2	3.1	1.7	2.9
6	32.7	36.1	26.4	20.5	31.2	27.7
7	3.2	6.7	3.5	9.1	1.8	4.5
8	7.7	11.5	10.7	7.6	13.6	7.7

* WT = Washington Team ** WSC = Washington Self-Contained
*** LSC = Longfellow Self-Contained

There were two types of ANOVA applied to the Flanders Model. The first exactly parallels the 2x3 factorial division shown above in Table 9 for the Withall Model analysis. In addition to this a 2x2 factorial ANOVA was computed on Teams only. Revised Category 5 (Criticizes or Justifies Authority) was omitted from these analyses due to its low frequency.

The results of the 2x3 ANOVA for each of the seven remaining Flanders Categories are shown below in Table 13 (Organizations and Grade Levels correspond to the division shown in Table 8).

TABLE 13

F-Ratios of Revised Flanders Categories by Category and Source

First Semester, 1961-62

Source	1	2	3	4	6	7	8
Organization (0)	1.52	8.07**	4.50*	-	3.13	1.07	6.73*
Grade Level (L)	1.86	-	3.43	-	4.82*	-	-
0 x L	5.65**	-	-	-	2.30	5.35**	1.73

$^{*}p < .05$
$^{**}p < .01$

Table 14, below, shows a breakdown of the mean percentage scores by organization, with the corresponding Duncan minimum ranges.

TABLE 14

Mean Percentage Scores of Revised Flanders Categories by Organization

First Semester, 1961-62

Category	Washington Team	Washington Self-Contained	Longfellow Self-Contained	Duncan Range
1	8.7	7.5	8.8	
2	12.6	15.3	18.5	3.1
3	21.0	10.8	15.0	6.9
4	15.0	14.0	14.1	
5	2.5	2.0	3.6	
6	26.6	33.6	27.1	
7	6.3	4.3	4.0	
8	7.6	12.5	9.2	3.1

The Duncan range shows the Team vs. Washington Self-Contained comparison to be non-significant for Category 2 (Asks Questions). But there is a significant difference in asking questions between Washington Team and Longfellow Self-Contained.

For Category 3 (Lectures), the same pattern is found as with the corresponding Withall Category 3: The Team was significantly higher than Washington Self-Contained, while the Washington Self-Contained vs. Longfellow Self-Contained difference is non-significant.

In Category 8 (Silence or Confusion), Washington Self-Contained was significantly higher than both other organizations.

The mean percentages by grade level, shown below in Table 15, reveal the cause of the significance of grade level in Category 6 (Student Response). The primary level is considerably higher than the intermediate level in this category.

TABLE 15

Mean Percentage Scores of Revised Flanders Categories by Grade Level

First Semester, 1961-62

	Grade Level	
Category	Primary	Intermediate
1	8.8	7.9
2	15.4	15.6
3	12.9	18.2
4	13.9	14.8
5	2.8	2.6
6	31.7	26.5
7	4.5	5.2
8	9.9	9.6

The interactions in Category 1 (Accepts Feeling, Praises, or Encourages, Accepts or Uses Ideas of Students) and Category 7 (Student Initiation) can be explained by reference to Table 12, above.

The interaction in Category 1 was caused primarily by the reversed high and low frequencies by grade level in Washington Team and

Self-Contained classes: The Primary Self-Contained classes were high in Category 1 while Intermediate Self-Contained were low, but the Primary Team classes were comparatively low in Category 1, while Intermediate Team classes were high. Longfellow Self-Contained classes followed the same pattern as Washington Self-Contained, but the differences were not as great and this may be another source of the interaction.

The major source of interaction in Category 7 was the low percentages for the Primary Team and the high percentages for the Intermediate Team, without consistent lows and highs for the self-contained classes.

Table 16, below, shows the breakdown of the Revised Flanders results for intern and master teachers within the Team.

TABLE 16

Mean Percentages of Revised Flanders Categories for the Team by Type of Teacher and Grade Level

First Semester, 1961-62

	Primary Level		Intermediate Level	
Category	Interns	Masters	Interns	Masters
1	8.8	6.5	8.0	11.1
2	14.7	12.4	12.0	11.3
3	9.5	28.1	13.5	32.5
4	13.0	16.2	17.5	13.4
5	2.0	1.5	3.7	2.8
6	38.0	27.5	26.0	14.8
7	4.2	2.3	11.1	7.2
8	9.8	5.5	8.2	6.9

An Analysis of Variance was computed on these frequencies, divided as shown below in Table 17. Master and intern teachers with the Primary and Intermediate Teams were compared for percentages of observations within Flanders categories.

TABLE 17

Division of 2x2 Analysis of Variance for Revised Flanders Categories

First Semester, 1961-62

Grade Level	Master Teachers	Interns
Primary Team		
Intermediate Team		

The results of the 2x2 factorial ANOVA are given below in Table 18.

TABLE 18

F-Ratios of Revised Flanders Categories
for Team Teachers by Category and Source

First Semester, 1961-62

	Category Number							
	1	2	3	4	5	6	7	8
Grade Level (L)	1.77	1.53	-	-	2.61	19.12**	9.22**	-
Type of Teacher (T)	-	-	11.65**	-	-	15.61**	2.06	2.10
L x T	3.55	-	-	-	1.20	-	-	-

$^{**}p < .01$

The sources of the grade level significances can be seen in Table 19.

TABLE 19

Mean Percentages of Revised Flanders Categories for the Team by Grade Level

First Semester, 1961-62

Category	Primary	Intermediate
1	7.6	9.5
2	13.5	11.7
3	18.8	23.1
4	14.6	15.5
5	1.8	3.1
6	32.7	20.5
7	3.3	9.1
8	7.7	7.5

The primary level is considerably higher than the intermediate level in Category 6 (Student Response) and considerably lower in Category 7 (Student Initiation).

Sources of type of teacher significance, on the other hand, can be seen in Table 20.

TABLE 20

Mean Percentages of Revised Flanders Categories for the Team by Type of Teacher

First Semester, 1961-62

Category	Interns	Masters
1	8.4	8.9
2	13.3	11.9
3	11.5	30.5
4	15.2	14.8
5	2.8	2.1
6	32.1	21.0
7	7.7	4.7
8	9.0	6.1

In Category 3 (Lectures) the master teachers were much higher than the interns. In Category 6 (Student Response) the interns were high.

2) Second Semester, 1961-62: Revised Withall: Frequency per Category

Tables 21-24, below, show the mean number of frequencies of Revised Withall categories for arithmetic, science, social studies, and reading by grade level which resulted from the second semester observations.

TABLE 21

Mean Number of Frequencies of Revised Withall Categories for Arithmetic by Grade Level

Second Semester, 1961-62

	Grade					
Category	1	2	3	4	5	6
1	4.8	6.2	5.2	7.0	5.2	6.0
2	7.2	2.2	6.5	4.2	8.2	6.5
3	6.5	8.5	4.5	10.2	8.0	5.8
4	8.0	6.8	7.5	6.2	5.0	2.8
5	7.0	9.5	9.5	5.8	8.0	14.2
6	2.5	2.8	2.8	2.5	1.5	0.7

TABLE 22

Mean Number of Frequencies of Revised Withall Categories for Science by Grade Level

Second Semester, 1961-62

	Grade					
Category	1	2	3	4	5	6
1	4.0	4.0	7.0	7.2	2.2	7.8
2	8.5	9.2	11.0	7.5	7.8	11.0
3	9.8	10.0	6.2	13.0	15.5	9.2
4	4.8	6.0	4.0	1.5	3.0	2.0
5	8.0	5.8	6.5	4.2	6.0	5.0
6	1.0	1.0	1.3	2.5	1.3	1.0

TABLE 23

Mean Number of Frequencies of Revised Withall Categories for Social Studies by Grade Level

Second Semester, 1961-62

	Grade					
Category	1	2	3	4	5	6
1	4.0	5.0	8.0	8.2	5.0	7.2
2	17.8	9.5	9.5	7.0	7.8	9.2
3	4.8	7.2	10.0	14.8	15.0	11.5
4	3.5	3.8	2.2	2.0	1.8	2.2
5	5.8	7.0	4.8	3.5	5.5	5.5
6	0.5	3.5	1.5	0.5	1.0	0.3

TABLE 24

Mean Number of Frequencies of Revised Withall Categories for Reading by Grade Level

Second Semester, 1961-62

	Grade					
Category	1	2	3	4	5	6
1	4.0	3.8	6.2	5.2	4.2	7.5
2	9.5	15.8	11.8	7.0	3.0	7.8
3	6.0	1.8	7.2	7.2	20.2	8.2
4	3.8	5.2	4.5	5.5	3.0	3.0
5	11.5	7.5	4.5	8.0	3.2	9.0
6	1.3	2.0	1.8	3.0	2.3	0.5

Analyses of variance were computed for each of the revised categories, while grade and subject area were treated as independent variables. The resulting division of each analysis was therefore 4x6 as shown below in Table 25. Since subject area was controlled in this study, frequencies were determined for each teacher in each subject area. (Grades were treated independently rather than as primary and intermediate level as in the previous study which included the teams).

TABLE 25

Division of 4x6 Analysis of Variance for Revised Withall Categories*

Second Semester, 1961-62

Subject	1st Grade	2nd Grade	3rd Grade	4th Grade	5th Grade	6th Grade
Arithmetic						
Science						
Social Studies						
Reading						

* This is not a factorial division since teachers are nested within grades, that is, a single teacher teaches only one grade.

In this analysis an error term was introduced for variation between teachers within grades, summed over all the grades; the degrees of freedom (df) were computed as shown below:

Source	df
Grade (g)	5
Teacher/Grade (t/g)*	6
Subject Area(s)	3
g x s	15
t/g and s**	18
Total	47

The results of the ANOVA's for the Revised Withall Categories are given in Table 26. (Revised Category 6 is omitted due to low frequency.)

TABLE 26

F-Ratios of Revised Withall Categories for Self-Contained Classes by Category and Source

Second Semester, 1961-62

	Category				
Source	1	2	3	4	5
Subject Area	-	3.79*	4.14*	12.43**	3.15*
Grades	1.04	1.04	1.29	-	-
s x g	-	1.77	2.44*	-	1.09

$^*p < .05$
$^{**}p < .01$

---

* The mean square associated with this term was used as the error term for testing g.

** The mean square associated with this term was used as the error term for testing s and g x s.

Four out of the five categories showed a significant effect for subject area. The effect of the four significant subject areas can be seen in Table 27, by comparing the variation of means of Categories 2-5 (significant) to the variation of Category 1 (non-significant).

TABLE 27

Mean Number of Frequencies of Revised Withall Categories by Subject Area and Category

Second Semester, 1961-62

	Subject Area			
Category	Arithmetic	Reading	Science	Social Studies
1	5.7	5.2	5.4	6.2
2	5.8	9.2	9.2	10.1
3	7.3	8.5	10.6	10.6
4	6.1	4.2	3.6	2.6
5	9.0	7.3	5.9	5.4

The significant interaction in Revised Category 3 (Gives Information, Gives Opinion, Gives Analysis) was caused by the extreme difference in frequency of this category for certain subject areas, particularly between grade two, which was very low, and grade five, which was very high; there was considerable variation in the remaining frequencies.

This can be seen in the distribution of means for Category 3 shown in Table 28.

TABLE 28

Mean Number of Frequencies of Revised Withall Category 3 by Subject Area and Grade Level Second Semester, 1961-62

	Subject Area			
Grade Level	Arithmetic	Reading	Science	Social Studies
1	6.5	6.0	9.8	4.8
2	8.5	1.8	10.0	7.2
3	4.5	7.2	6.2	10.0
4	10.2	7.2	13.0	14.8
5	8.0	20.2	15.5	15.0
6	5.8	8.2	9.2	11.5

3) First Semester, 1962-63

a) Flanders: Frequency per Category

The mean number of frequencies of the Flanders Categories for arithmetic, social studies and reading by organization are given in Tables 29-31.

TABLE 29

Mean Number of Frequencies of Flanders Categories for Arithmetic by Organization

First Semester, 1962-63

Category	Washington Team	Washington Self-Contained	Longfellow Self-Contained
1	1.1	1.0	1.3
2	2.9	1.3	2.3
3	49.1	39.8	37.5
4	47.8	66.4	59.0
5	37.3	29.0	21.6
6	91.0	88.1	100.5
7	12.0	5.6	6.5
8	89.8	125.0	116.0
9	11.9	11.1	9.5
10	56.3	32.8	44.9

TABLE 30

Mean Number of Frequencies of Flanders Categories for Social Studies by Organization

First Semester, 1962-63

Category	Washington Team	Washington Self-Contained	Longfellow Self-Contained
1	1.3	.6	1.8
2	5.3	4.3	7.8
3	28.8	52.5	33.6
4	55.8	60.8	68.9
5	77.4	75.0	87.6
6	58.0	37.5	45.6
7	17.1	9.4	7.1
8	103.6	104.3	110.5
9	34.1	34.1	22.3
10	17.8	20.6	13.9

TABLE 31

Mean Number of Frequencies of Flanders Categories for Reading by Organization

First Semester, 1962-63

Category	Washington Team	Washington Self-Contained	Longfellow Self-Contained
1	4.9	2.2	.5
2	12.4	3.9	1.0
3	38.6	44.1	53.6
4	42.8	54.9	72.4
5	33.5	47.6	53.6
6	85.0	60.8	35.0
7	12.4	11.3	7.5
8	123.6	126.6	123.8
9	21.6	17.9	11.0
10	24.3	29.8	40.8

A 3x3 factorial ANOVA was computed for each of the ten categories of the Flanders Model: Three subject areas, arithmetic, social studies, and reading were treated as shown in Table 32.

TABLE 32

Division of 3x3 Analysis of Variance for Flanders Categories

First Semester, 1962-63

Subject Area	Washington Team	Washington Self-Contained	Longfellow Self-Contained
Arithmetic			
Social Studies			
Reading			

The results of the ten ANOVA's for Flanders Categories is shown below in Table 33.

TABLE 33

F-Ratios of Flanders Categories by Category and Source

First Semester, 1962-63

Category	Organization	Subject	TxS
1	1.60	2.14	2.11
2	1.81	1.98	2.02
3	-	-	2.75*
4	2.25	-	-
5	-	4.32	-
6	1.23	7.40**	1.07
7	4.06*	-	-
8	-	-	-
9	2.01	10.05**	-
10	-	5.41**	-

$^{*}p < .05$
$^{**}p < .01$

Organization was significant in only one instance, Category 7, (Criticizes or Justifies Authority). The self-contained classes were close in frequency while the team was significantly higher according to the Duncan procedure. These differences can be seen in the breakdown of mean frequencies by organization shown below in Table 34.

The significant interaction of Category 3 (Accepts or Uses Ideas of Students) was due to the crossover between the two self-contained classroom organizations, shown below in Table 35. Washington Self-Contained was high in Category 3 interactions in social studies, and low in reading, while Longfellow Self-Contained was low in social studies and high in reading.

TABLE 34

Mean Number of Frequencies of Flanders Categories by Organization

First Semester, 1962-63

Category	Washington Team	Washington Self-Contained	Longfellow Self-Contained	Duncan Range
1	2.4	1.3	1.2	
2	6.8	3.1	3.7	
3	38.8	45.4	41.6	
4	48.8	60.7	66.8	
5	49.4	50.2	54.3	
6	78.0	62.1	60.4	
7	13.8	8.8	7.0	5.0
8	105.7	118.6	116.8	
9	22.5	21.0	14.3	
10	32.8	27.7	33.3	

TABLE 35

Mean Number of Frequencies of Flanders Category 3 by Organization and Subject Area

First Semester, 1962-63

Subject	Washington Team	Washington Self-Contained	Longfellow Self-Contained
Social Studies	28.8	52.5	33.6
Reading	38.6	44.1	53.6
Arithmetic	49.1	39.8	37.5

V. DESCRIPTIVE ANALYSIS

A. Method of Analysis

The descriptive method of analyzing interaction data suggested by Flanders (55) was used on the matrix tabulations obtained from observations in the third study, (first semester 1962-63). This consists of the summary tabulation of specific areas of the data matrix. These areas were identified by Flanders and are summarized below in Table 36.

There are two modifications of the Flanders areas. Flanders describes the 6-7, 7-7, 7-6, 6-6 combination of cells as the "vicious cycle" of "giving directions" followed by "criticism" and vice versa. However, it was felt that 6-6, which indicates "giving directions" followed by "giving directions" does not necessarily fall in with "criticism," and in fact, should not be definitely associated with Flanders "vicious cycle". To locate disciplinary interaction more precisely, the column of sevens and 7-6 were examined. This area was named "Disciplines Students" and does not include the 6-6 interaction cell.

Flanders includes Category 7 (Criticizes or Justifies Authority) in Ignores Student Participation. It has been omitted here in order that Disciplines Students and Ignores Student Participation will be mutually exclusive areas of the matrix, and may be viewed as separate aspects of Direct Teacher Influence.

There were no statistical tests of the between group differences. The assumptions for the usual parametric and non-parametric tests do not hold due to the unique interdependence of the matrix elements. The discussion in the descriptive section which follows is based entirely upon inspection of tabulations.

TABLE 36

Interaction Areas of the Flanders Matrix

Interaction Area	Matrix Area	Description
Indirect Teacher Influence	Columns 1, 2, 3, 4	Teacher Accepts Feeling Praises or Encourages Accepts or Uses Ideas of Student Asks Questions
Direct Teacher Influence	Columns 5, 6, 7	Teacher Lectures Gives Directions Criticizes or Justifies Authority
Student Talk	Columns 8, 9	Student Response or Initiation
Silence or Confusion	Column 10	
Students Develop Own Ideas	Cells 9-9, 8-8	Student Talk followed by continued Student Talk
Constructive Integration	Column 1, 2 ,3- Row 1, 2, 3	Indirect Teacher Influence (except Asks Questions) followed by Indirect Teacher Influence
Disciplines Students	Column 7, Cell 7-6	Criticizes or Justifies Authority and Gives Directions followed by Criticism
Encourages Student Participation	Columns 8, 9- Rows 1, 2, 3, 4	Student Talk followed by Indirect Teacher Influence
Ignores Student Participation	Columns 8, 9- Rows 5, 6	Student Talk followed by Direct Teacher Influence except Criticism
Content Cross	Columns 4, 5- Rows 4, 5*	All interaction either preceded by or following Asks Questions and Lectures

* Cells common to both are counted only once.

B. Results

Table 37, below, shows percentages of total observations on each organization which correspond to each interaction area. The percentages may therefore be read as the average amounts of time in a class period occupied by student-teacher interaction of the types described.

TABLE 37

Percentages of Total Observations of Flanders Interaction Areas by Organization

First Semester, 1962-63

Interaction	Washington Team	Washington Self-Contained	Longfellow Self-Contained
Indirect Teacher Influence	24.3	27.7	28.4
Encourages Student Participation	8.8	11.2	11.4
Constructive Integration	3.2	3.2	3.2
Direct Teacher Influence	35.4	30.3	30.5
Ignores Student Participation	4.1	3.3	3.8
Disciplines Students	3.9	2.4	2.0
Student Talk	32.1	35.0	32.8
Students Develop Own Ideas	15.7	18.5	15.9
Silence or Confusion	8.2	7.0	8.3
Content Cross	34.3	39.4	43.0

It will be noted that four interaction areas are mutually exclusive and include all interaction within the class period: Indirect Teacher Influence; Direct Teacher Influence; Student Talk; and, Silence or Confusion. The Team is clearly higher in Direct Teacher Influence and lower in Indirect Teacher Influence than either self-contained organization. In Student Talk and Silence or Confusion the organizations are approximately equal. *

As a part of Direct Teacher Influence, ** Disciplines Students is disproportionately higher in the Team than self-contained classes, compared to the proportion of total Direct Teacher Influence. This disproportion was reflected by the fact that only Category 7 (Criticizes or Justifies Authority) proved significant in the Analysis of Variance across organizations (See Table 33 above).

Content Cross was somewhat lower in the Team than in the self-contained classes, indicating that the teachers of self-contained classes occupied more of the total teaching time with lecturing and asking questions, than the team teachers. (Note that total teaching time is almost the same in the three organizations, being the sum of Direct and Indirect Teacher Influence.)

---

* See pages 15-17 for mention of the other studies on the effect of team organization on student discussion.

** Greater discipline problems under the team organization were not reported in almost all subjective evaluations of other team projects (pp. 15-17). Few of these projects used interns.

TABLE 38

Percentages of Total Observations of Flanders Interaction Areas by Organization and Grade Level

First Semester, 1962-63

Interaction	Washington Team		Washington Self-Contained		Longfellow Self-Contained	
	Prim.	Int.	Prim.	Int.	Prim.	Int.
Indirect Teacher Influence	27.9	20.7	27.8	27.7	31.2	25.5
Encourages Student Participation	10.9	6.8	12.0	10.3	12.1	10.7
Constructive Integration	3.6	2.8	3.7	2.8	3.4	3.0
Direct Teacher Influence	32.9	37.9	26.3	36.4	28.2	32.8
Ignores Student Participation	4.7	3.6	3.8	2.8	3.7	4.0
Disciplines Students	3.1	4.7	1.9	2.9	3.2	.9
Student Talk	35.7	28.6	39.4	30.6	30.5	35.2
Students Develop Own Ideas	17.1	14.2	21.4	15.6	12.3	19.4
Silence or Confusion	3.5	12.9	6.6	7.4	10.1	6.5
Content Cross	39.8	28.9	32.9	46.0	38.5	47.5

In Table 38, the percentages for interaction areas have been broken down into primary and intermediate levels within each organization. Although the differences between grade levels are fairly consistent for the three organizations, and not generally very great, there are some notable inconsistencies: Figures 4-8 present graphs showing the changes in interaction area percentages of total observations from primary to intermediate level in the three organizations.

In overall teacher influence the trends between grade levels are consistent: Indirect Teacher Influence is higher in primary levels than intermediate levels, while Direct Teacher Influence is higher in intermediate levels than primary levels.

A crossover is apparent in Student Talk (Figure 6) between Longfellow Self-Contained and the two organizations within Washington School.

Figure 7 reveals the first difference between grade levels which separates the Team from the self-contained organizations. While Silence or Confusion dropped off 3.6% in Longfellow Self-Contained and increased 0.8% in Washington Self-Contained, from the primary to intermediate levels, it increased 9.4% between the Team levels, such that Silence or Confusion was markedly higher in the Intermediate Team.

In Content Cross, the difference between team levels is again markedly different from the difference between self-contained levels. Here it is apparent that the low Content Cross percentage of the Team was contributed entirely by the Intermediate Team, while the Primary Team was higher in Content Cross than the primary levels of the self-contained organizations.

In further exploring the apparent differences between the Primary and Intermediate Teams, interaction area percentages were broken down for interns and master teachers within the Team. This breakdown was restricted

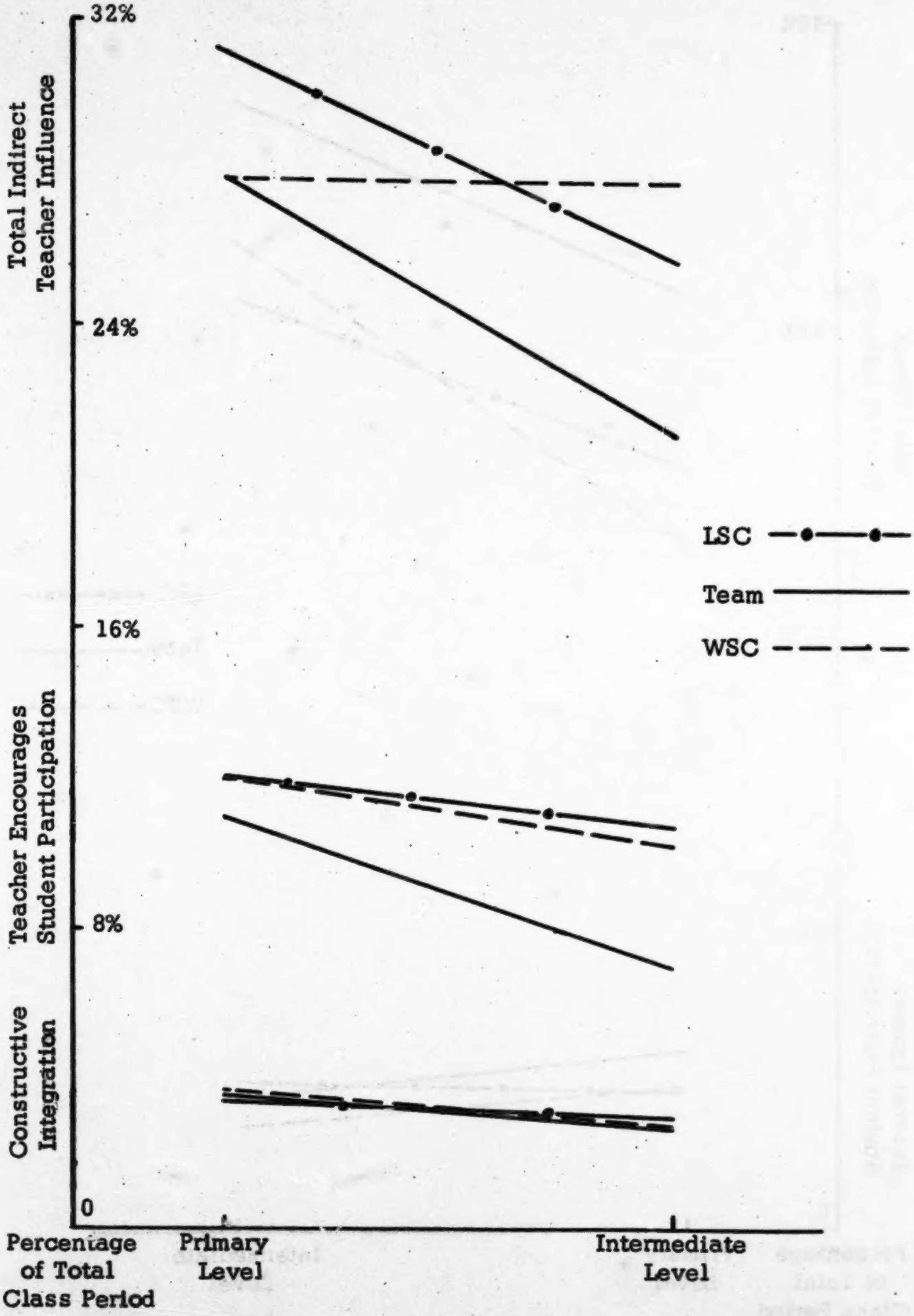

Figure 4: Indirect teacher influence by organization and grade level.

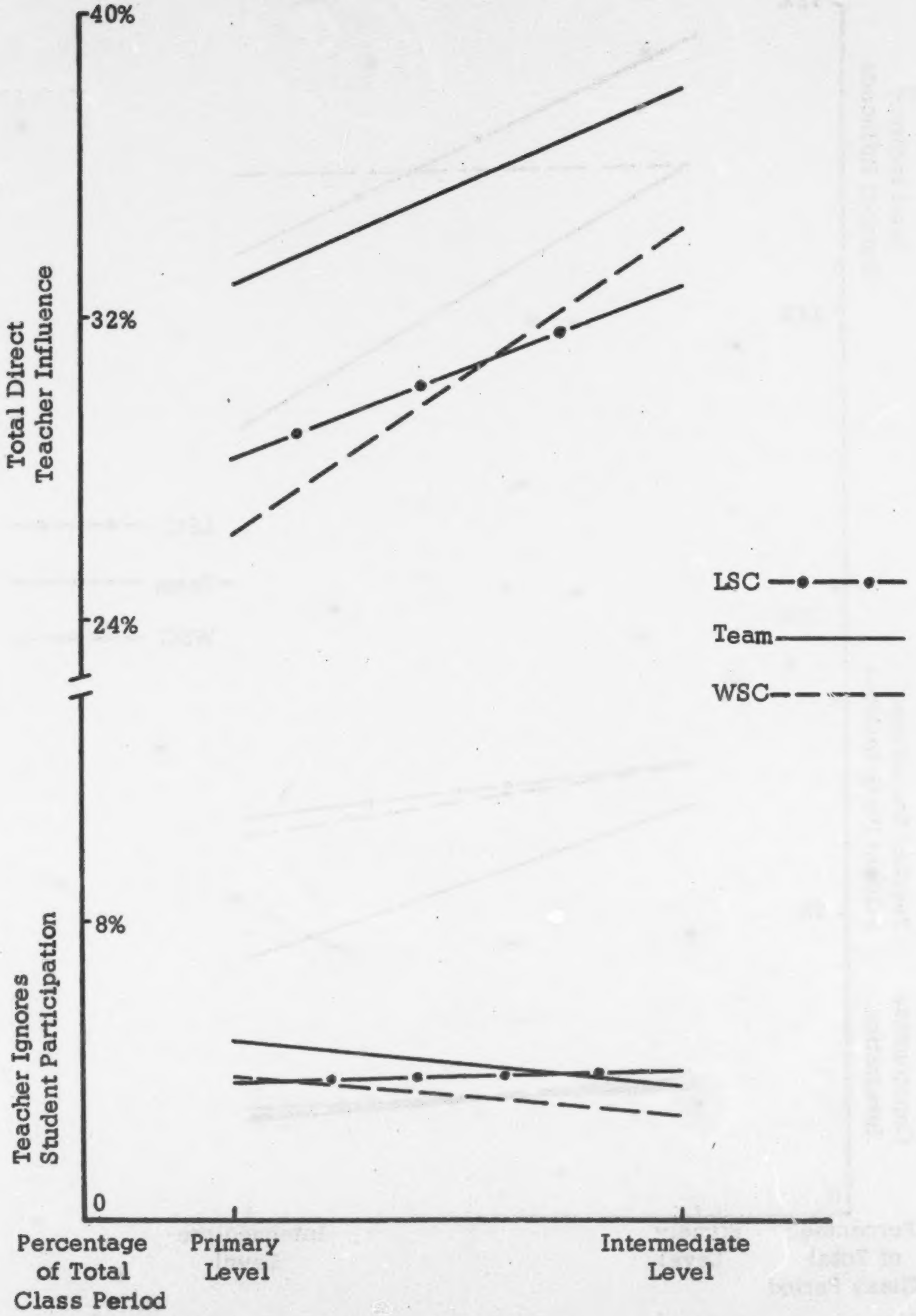

Figure 5: Direct teacher influence by organization and grade level.

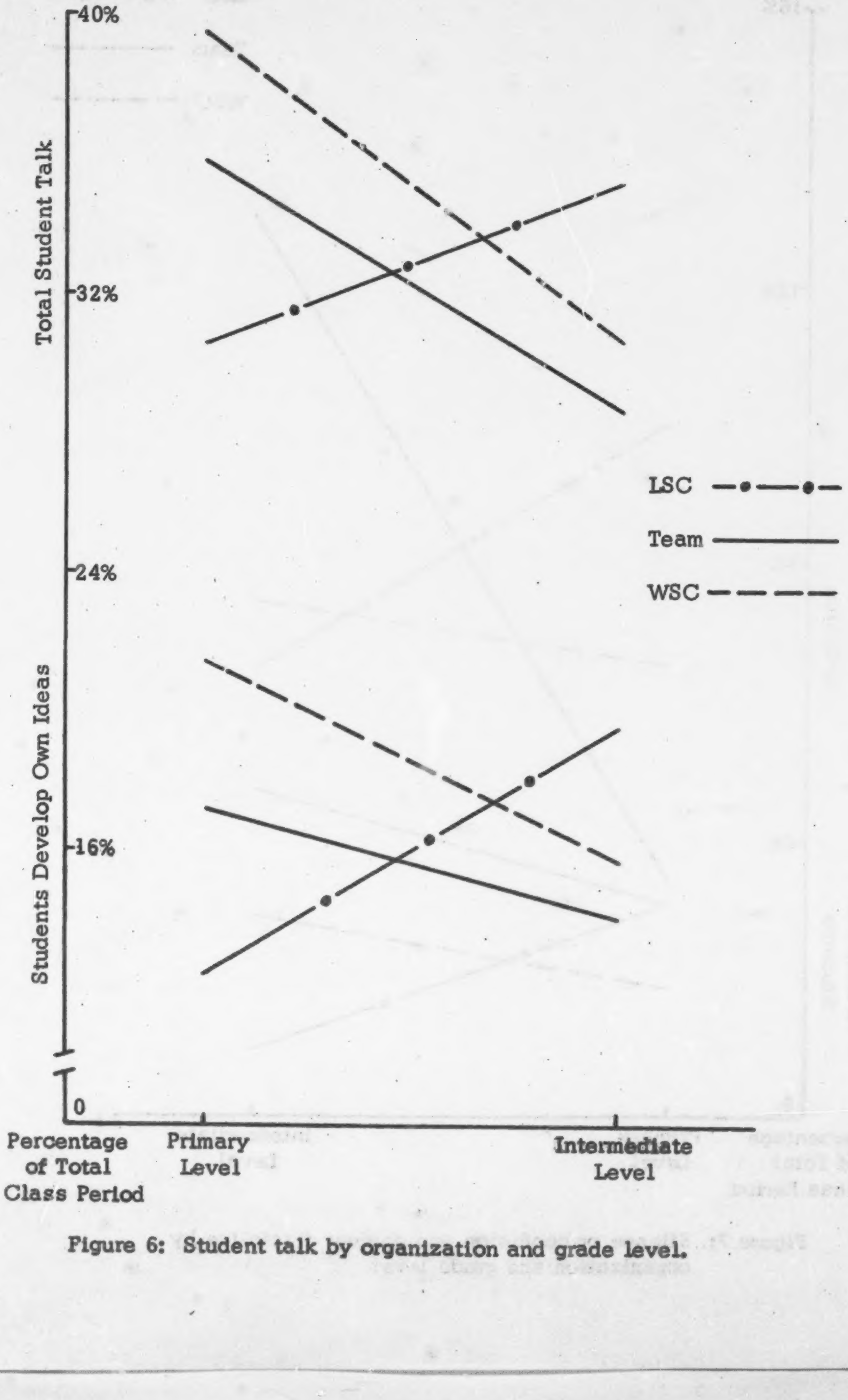

Figure 6: Student talk by organization and grade level.

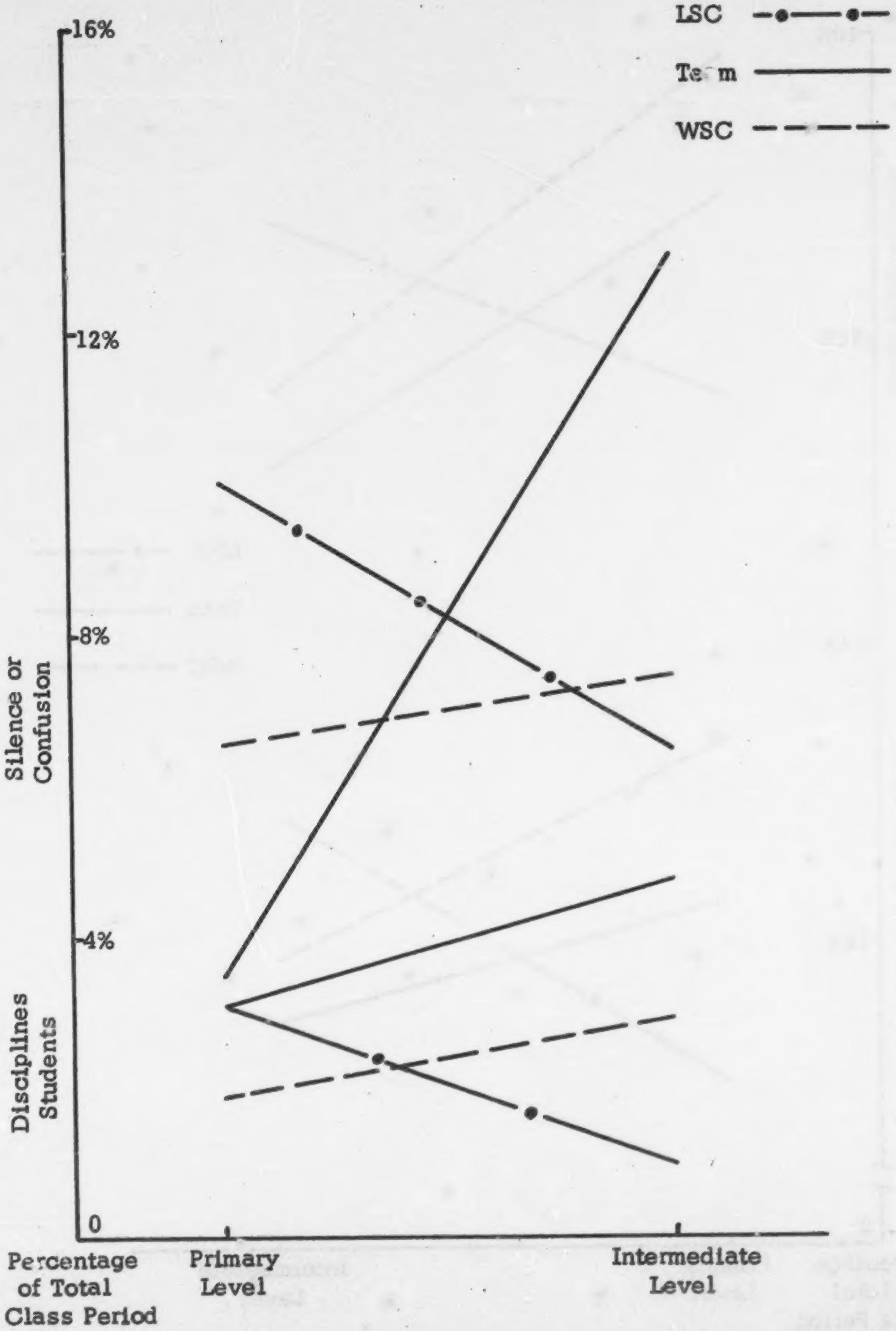

Figure 7: Silence or confusion and student discipline by organization and grade level

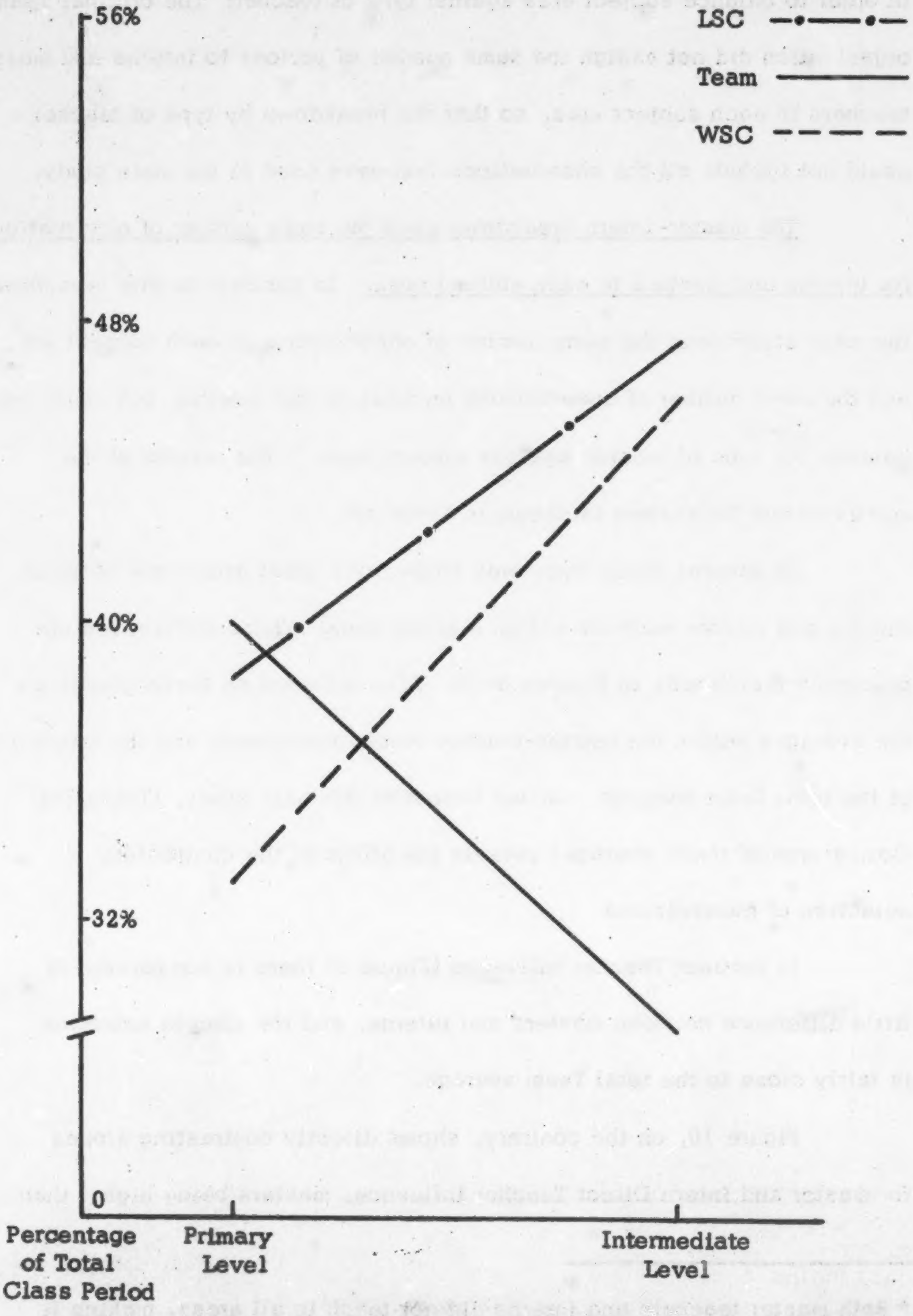

Figure 8: Content cross by organization and grade level.

in order to balance subject area against type of teacher: The original Team organization did not assign the same number of periods to interns and master teachers in each subject area, so that the breakdown by type of teacher could not include all the observations that were used in the main study.

The master-intern breakdown used the same number of observations for interns and masters in each subject area. In contrast to this breakdown, the main study used the same number of observations in each subject area and the same number of observations on masters and interns, but could not balance the type of teacher against subject area. * The results of the master-intern breakdown is shown in Table 39.

In several cases there was evidently a great difference between interns and master teachers within a given team. These differences are presented graphically in Figures 9-13. Also included on these graphs are the averages within the master-teacher breakdown sample and the averages of the total Team changes, carried over from the main study, (Table 38). Comparison of these averages reveals the effect of the contrasting selection of observations.

In Indirect Teacher Influence (Figure 9) there is comparatively little difference between masters and interns, and the sample selection is fairly close to the total Team average.

Figure 10, on the contrary, shows directly contrasting slopes for master and intern Direct Teacher Influence, masters being higher than

---

* Both master teachers and interns did not teach in all areas, making it impossible to obtain this balance in the main study.

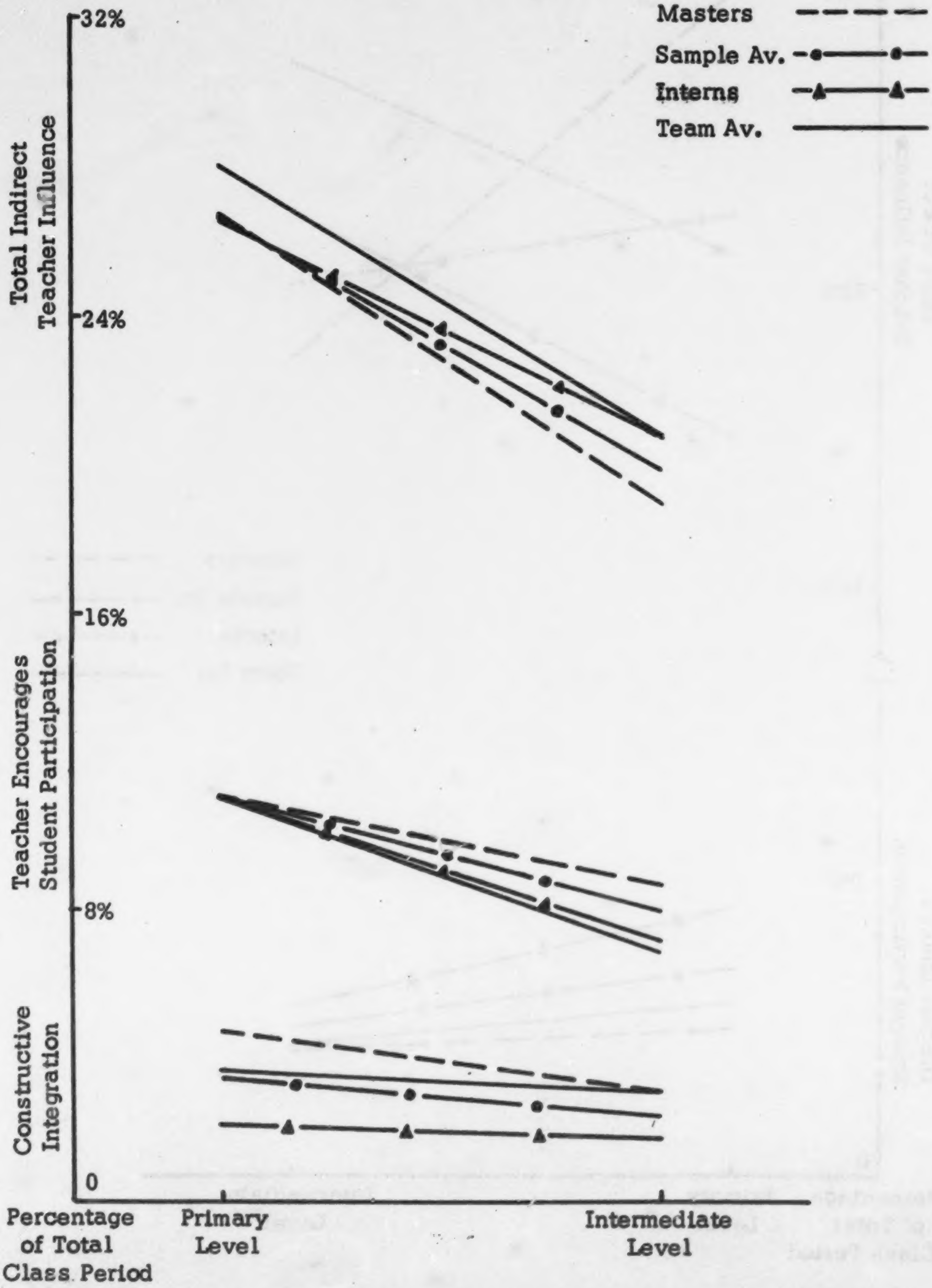

Figure 9: Indirect teacher influence by type of teacher and grade level.

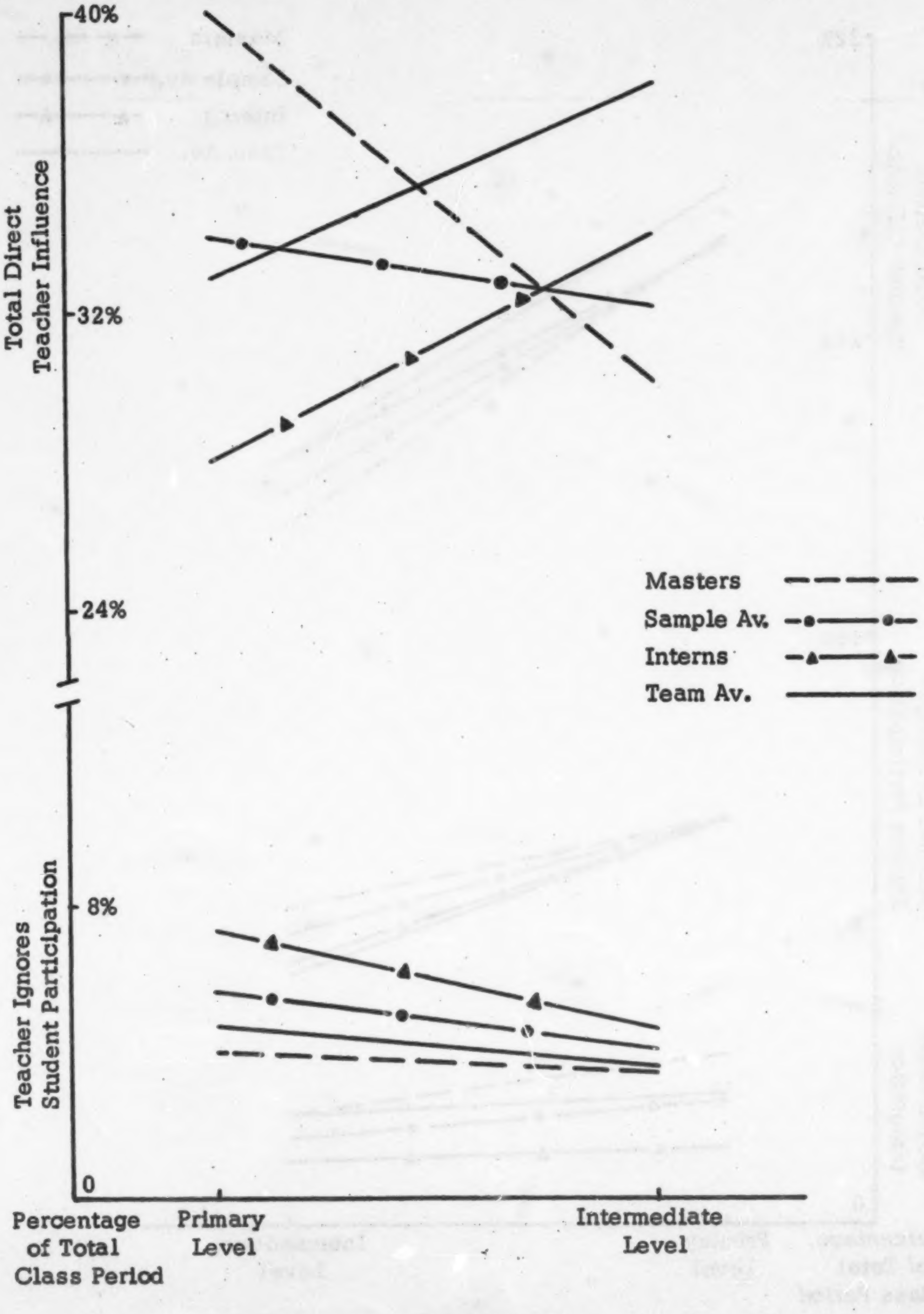

Figure 10: Direct teacher influence by type of teacher and grade level.

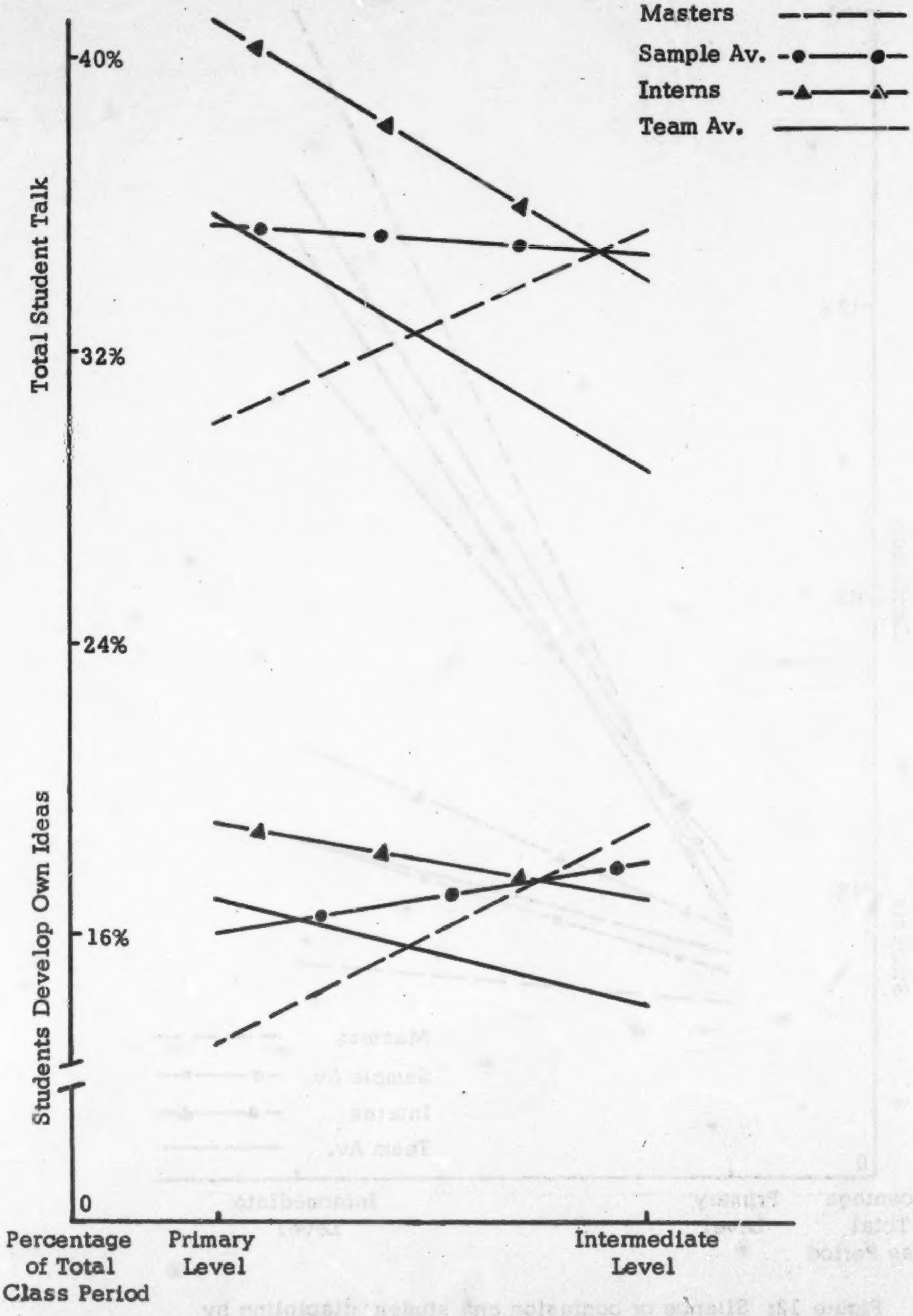

Figure 11: Student talk by type of teacher and grade level.

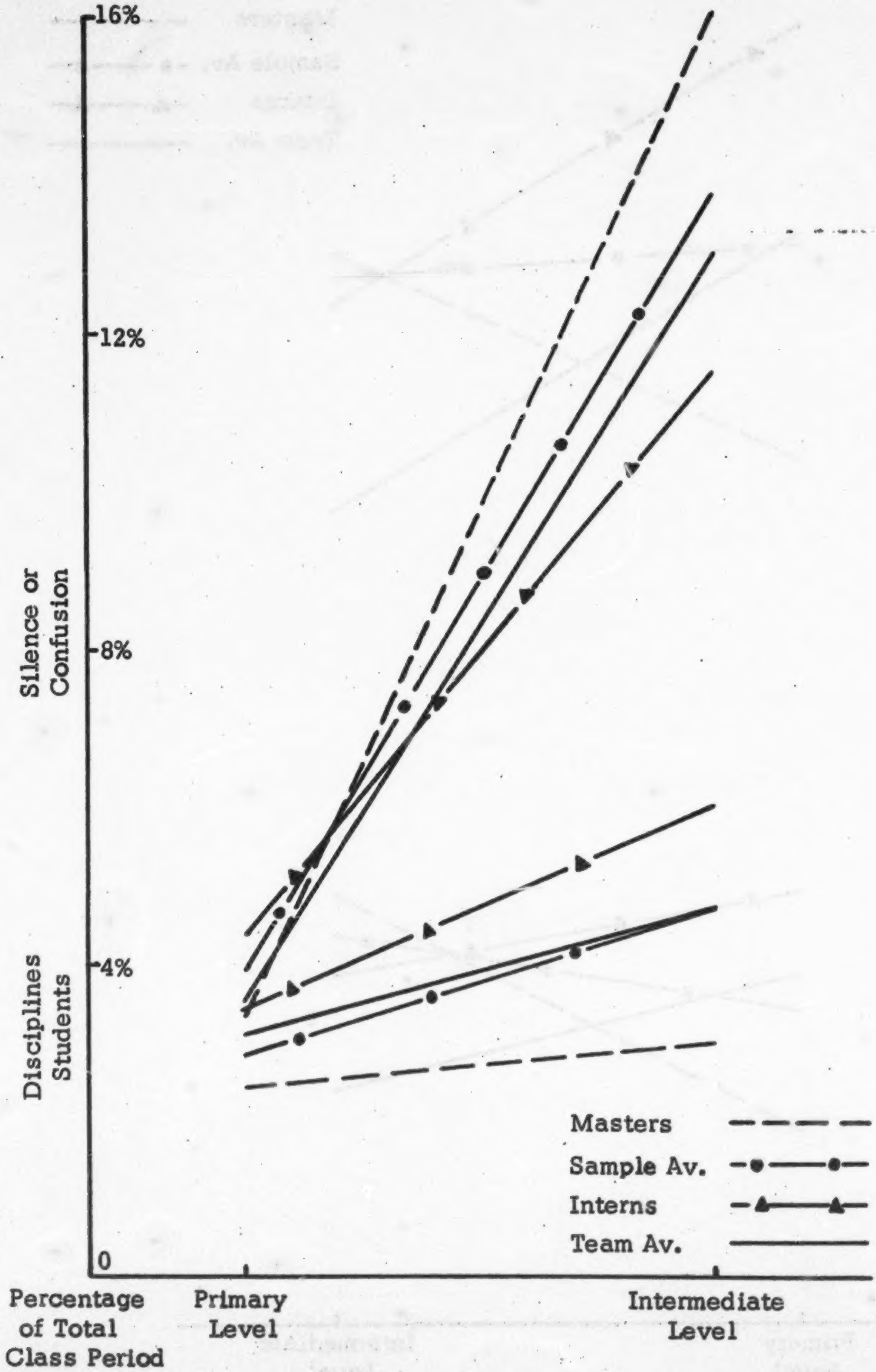

Figure 12: Silence or confusion and student discipline by type of teacher and grade level.

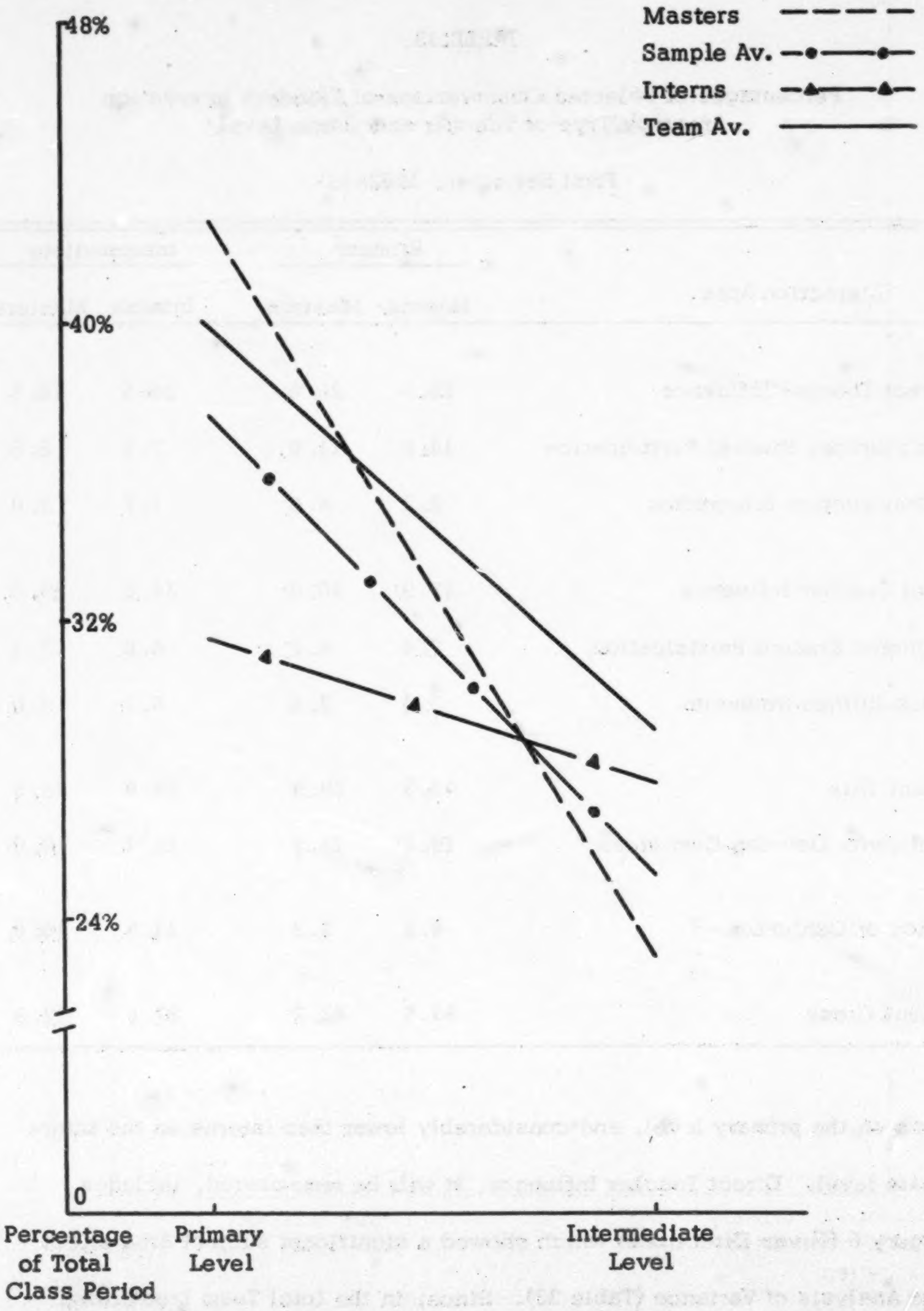

Figure 13: Content cross by type of teacher and grade level.

TABLE 39

Percentages of Selected Observations of Flanders Interaction Areas by Type of Teacher and Grade Level

First Semester, 1962-63

Interaction Area	Primary		Intermediate	
	Interns	Masters	Interns	Masters
Indirect Teacher Influence	26.5	26.6	20.6	18.9
Encourages Student Participation	11.0	11.0	7.1	8.6
Constructive Integration	2.1	4.6	1.7	3.0
Direct Teacher Influence	27.9	40.0	34.0	29.8
Ignores Student Participation	7.4	4.0	4.6	3.3
Disciplines Students	3.4	2.4	6.1	3.0
Student Talk	41.3	29.9	33.9	35.3
Students Develop Own Ideas	19.0	13.2	17.1	19.0
Silence or Confusion	4.3	3.3	11.5	16.0
Content Cross	31.5	42.7	27.6	22.9

interns on the primary level, and considerably lower than interns on the intermediate level. Direct Teacher Influence, it will be remembered, includes Category 6 (Gives Directions) which showed a significant subject area effect in the Analysis of Variance (Table 33). Since, in the total Team breakdown, of Table 33, subject matter is confounded with type of teacher, the difference in slopes between the average of the sample and the total Team average may be partially accounted for by effect of subject area. This effect may be acting again on the averages of Student Talk (Figure 11), remembering that Category 9 (Student Initiated Response) showed subject area significance.

In Student Talk, however, the difference between masters and interns is large on the primary level, 29.9% compared to 41.3%, while this difference is almost erased on the intermediate level.

Figure 12 shows that in the areas of Silence or Confusion and Student Discipline, the sample averages are quite close to the totals, and the differences between masters and interns in the Primary Team are not great. Interestingly, the master teachers in the Intermediate Team seem to have produced considerably more silence or confusion, while having less disciplinary problems.

In Content Cross, the difference between masters and interns on the primary level is very great, 42.7% as compared to 31.5% of total class time. On the Intermediate Level a smaller difference exists and in the reverse direction: 27.6% for the interns and 22.9% for the masters.

VI. CONCLUSIONS

In general, the interaction model analyses located less differences between classroom interaction in the team and self-contained organizations then might be expected. The first and third study results, both of which included all organizations, showed significant differences between the Team and both self-contained organizations in only two categories, as can be seen in the summary of significant organizational effects shown below in Table 40. These categories were Revised Flanders Category 2 and original Flanders Category 7.

In the first year, team teachers asked questions significantly less often than teachers in the self-contained organizations; and in the second year, team teachers criticized the students and/or justified their authority significantly more often than teachers in the self-contained classrooms. The former difference was not replicated in the second year study in which normal Flanders Category 4 (Asks Questions) did not show significance. The difference between the results of the first and second year studies in the category of criticism and justification of authority may be explained by the revision of categories in the first year study (see pages 55-56). In the first year when mild criticism was separated from strong criticism, mild criticism showed significant organizational effect, (Revised Flanders Category 3).

Comparison of Revised Withall and Flanders Categories 3, in Table 40, which may be said to show direct teacher influence, indicates a consistent pattern between the three organizations: The Team is higher than both self-

TABLE 40

Summary of Significantly Different Interaction Categories by Organization

1961-63

First Semester, 1961-62	Washington Team	Washington Self-Contained	Longfellow Self-Contained
Revised Withall Category Asks for Information Asks for Opinion or Analysis Seeks or Accepts Direction	low	low	high
Revised Withall Category 2 Listens	low	high	low
Revised Withall Category 3 Gives Information Gives Opinion Gives Analysis	high	low	intermediate
Revised Flanders Category 2 Asks Questions	low	high	high
Revised Flanders Category 3 Lectures Criticizes or Justifies Authority (strong criticism omitted)	high	low	intermediate
Revised Flanders Category 8 Silence or Confusion	low	high	low
Third Semester, 1962-63			
Flanders Category 7 Criticizes or Justifies Authority (strong criticism included)	high	low	low

contained organizations but only significantly higher than Washington Self-Contained classes. This effect would not have been as sharply discriminated in the third semester statistical analysis because the normal Flanders categorization subdivides types of teacher influence. Nevertheless, the descriptive analysis of the matrix tabulations revealed that the Team was again higher than self-contained organizations in the direct teacher influence area and lower in the indirect teacher influence area.

The three remaining categories which showed significant organizational effects, (Revised Withall Categories 1 and 2, Revised Flanders Category 8), did not show main differences between team and self-contained organizations, but rather pointed to significant differences between the control organizations, themselves.

Differences in student-teacher interaction patterns between grade levels proved to be less common than the differences between organizations. In the first semester study, three categories showed grade level significances. These are summarized in Table 41.

TABLE 41

Summary of Significantly Different Interaction Categories
by Grade Level

First Semester, 1961-62

	Primary	Intermediate
Revised Withall Category 3 Gives Information Gives Opinion Gives Analysis	low	high
Revised Withall Category 4 Gives Suggestions Gives Directions	high	low
Revised Flanders Category 6 Student Response	high	low

These results are perhaps to be expected, considering that lecturing of the type covered in Revised Withall Category 3 is normally higher in the upper grades, while direct suggestion and direction is normally higher in lower grades, and consequently, students usually do more talking (Revised Flanders Category 6) in lower grades and more listening in higher. There were indications, however, that these differences were greater between the Primary and Intermediate Teams than between primary and intermediate levels in self-contained classes.

In the second semester study, when grade levels were broken down into individual grades and only self-contained classes were observed, there were no significant effects of grade level whatever.

The descriptive analysis of the Flanders Matrix tabulations revealed that in several cases student-teacher interaction varied more between the Primary and Intermediate Teams than between the corresponding levels of the self-contained organizations. This difference was particularly evident in the area Content Cross, in which the Intermediate Team was considerably lower than the Primary Team while self-contained intermediate levels were considerably higher than their corresponding primary levels.

The descriptive breakdown of the Team by type of teacher and grade level in the third study can be compared to the ANOVA of type of teacher by grade level in the first study, Table 18 (p. 70). This study showed that Revised Flanders Categories 3 (Lectures, Criticizes and Justifies Authority) and 6 (Student Response) were significant for type of teacher. Table 42 shows the breakdown of this first study by grade level and type of teacher in comparison with the third study.

TABLE 42

Mean Percentages of Flanders Categories for the Team
by Type of Teacher, Grade Level, and Category
Comparing the Studies of the First Semester 1961-62
and the First Semester 1962-63

		School Year			
		1961-62		1962-63	
Grade Level	Interaction Area	Interns	Masters	Interns	Masters
Primary	Direct Teacher Influence	24.5	45.8	27.9	40.2
	Student Talk	42.2	29.8	41.3	29.9
Intermediate	Direct Teacher Influence	34.7	48.7	34.0	29.8
	Student Talk	37.1	22.0	33.9	35.3

In Table 42 Revised Category 3 is combined with Revised Category 4 (Gives Directions), and Revised Category 5 (Strong Criticism) for the sake of comparison to Direct Teacher Influence. The differences between masters and interns are clearly parallel in the Primary Team for both years, while there is a direct reversal in the Intermediate Team. In the first year, Intermediate Team masters used more Direct Teacher Influence than interns, while in the second year, they used less, and the combined Direct Teacher Influence was lower.

Averages on the Primary Level were 24.5% and 27.9% for interns in the first and second year, 45.8% and 40.2% for masters. On the Intermediate Level they were 34.7% and 34.0% for the interns, 48.7% and 29.8%, a considerable drop, for the masters.

Revised Category 6 (Student Response) is combined with Revised Category 7 (Student Initiation) in Table 42 for comparison with Student Talk in the descriptive analysis. This shows an almost exact replication of the first year results in the second year for the Primary Team: 42.2% and 41.3% respectively for the interns, and 29.8% and 29.9% respectively for the master teachers. In the Intermediate Team the trend is again reversed and the masters of the Intermediate Team of the second year elicited substantially more student talk than those of the Intermediate Team of the first year: These averages were 37.1% and 33.9% for the interns, first and second year, 22.0% and 35.3%, showing a considerable increase, for the masters.

The changes in the Intermediate Team may be summarized: While interns showed approximately the same amount of direct teacher influence and elicitation of student talk in both years of the study, master teachers showed high direct teacher influence the first year, and low direct teacher influence the second year, and they elicited low student talk the first year, and high student talk the second year.

Remembering that in Content Cross, the Intermediate Team masters were considerably lower than Primary Team masters in the second year, one notices a clear pattern emerging from these comparisons: the Intermediate Team fell consistently below expectations in the second year of the study, and not simply as a result of new interns, but primarily as the result of new master teachers: There was lower emphasis on content, and less direct teacher influence, and greater student talk. These results indicate great instability in the Intermediate Team, particularly considering the very close similarity of student-teacher interaction in the Primary Team for the first and second year.

Such observations clearly reflect the organizational difficulties which were discussed in the introductory chapter, and indicate a high degree of sensitivity of the interaction model to the team situation. These results indicate the importance that a team be led by teachers experienced in team teaching.

PART THREE

STUDENT ACHIEVEMENT AND PERSONALITY

PRECEDING PAGE BLANK-NOT FILMED

I. ACHIEVEMENT AND PERSONALITY: INTRODUCTION

Student personality and achievement were measured independently in the 1961-62 and 1962-63 academic years: In the first academic year, personality and achievement pre-tests were given in October, 1961, and post-tests in May, 1962. Only post-tests were administered in the second academic year. Personality post-tests were given in April, 1963, and achievement post-tests, in May, 1963. This schedule is shown, below:

SCHEDULE TWO

Personality and Achievement Studies
1961-1963

	First Year		Second Year
	Pre-Test	Post-Test	Post-Test
Personality	Oct. 1961	May 1962	April 1963
Achievement	Oct. 1961	May 1962	May 1963

ACHIEVEMENT AND PERSONALITY PROJECT
1961-1963

A. Academic Year 1961-62

The basic question raised in this study was the effect of organizational structure on student achievement and personality, which was measured by standardized achievement and personality tests. Sex and grade were included in the experimental design in order to isolate organizational effect from the possible effects of these variables.

An Analysis of Covariance, using achievement and personality pretests as covariates, was undertaken in order to control for initial differences between the students.* In the achievement study, I.Q. measurements on the students were also covaried out.** Thus, this study isolated organizational effect from sex and grade, and from possible initial differences between students.

An unequal cell analysis was used since the numbers in the various breakdowns of the design were not the same. It did not seem desirable to balance the number of students in the cells by random elimination of excess students since this would have involved discarding a considerable amount of the data. Separate analyses were computed for the various achievement and personality scores. The sources of organizational significance were investigated by the Duncan Range Test.

However, since the cell frequencies for the achievement and personality tests were unequal and, in many cases, quite different, a modification of the Duncan procedure was necessary. The procedure used was that outlined by Kramer (95). (The Duncan Range was computed on gain scores.) Three minimum gaps were computed when the organizations were compared. The gaps were between the three organizations in combinations of two. These three gaps are reported in the tables, as in Table 50, below. This procedure was used for the entire achievement and personality project.

---

* The Analysis of Covariance is described above, pp. 27-28.

** I.Q. had been measured before the study. (See pp. 33-34).

Grade, in this analysis, refers to the actual grades 1-6 rather than the grade levels, primary and intermediate. The first grade was deleted from the achievement statistical analysis since no I.Q. scores were available for these students.

B. Academic Year 1962-63

Several observations during the first year of study altered the procedure used in the following year: The achievement test which had been used did not appear to be properly articulated between levels, and there were also indications that the students were becoming overfamiliar with the test. Different achievement tests were therefore used for grades 3-6 inclusive, which are more reliable because of their increased length. (A different and better test for the first two grades, however, was not available at this time.)

Further, it was observed that the use of pre-tests caused a diversity of teacher reactions. Some teachers interpreted their students' achievements on these tests as a reflection of their personal capabilities, and having been cued by the pre-tests, they became biased towards the tests. Nevertheless, other teachers ignored the pre-tests. In order to eliminate this bias, the pre-tests were dropped in the second year study.

The I.Q. score was eliminated as a covariate since it was found that there were no significant differences across the cells in the grades for which I.Q. scores were available.

With the elimination of the covariates the analysis was changed to an Analysis of Variance for unequal cells. Grade and sex were again treated as independent variables.

In addition to the statistical analysis, a descriptive analysis of student achievement was undertaken in the second year study. This consisted primarily of a comparison of relative positions across the three organizations.

II. PERSONALITY: MEASUREMENT

The California Test of Personality (1953 Revision) by L. P. Thorpe, W. W. Clark, and E. W. Tiegs was used both years. Two levels of the test were used, primary for grades 1-3, and elementary for grades 4-6. In the first year, different forms of the test on each level were used for pre- and post-test. The test level and form corresponding to each grade is shown in Table 43.

TABLE 43

Personality Test Level by Grade, and Date
1961-63

Grade	Level Taken as Pre-Test, 1961	Level Taken as Post-Test, 1962	Level Taken as Post-Test, 1963
1 2 3	Primary Form BB	Primary Form AA	Primary Form BB
4 5 6	Elementary Form BB	Elementary Form AA	Elementary Form AA

These tests give two independent scores, personal adjustment and social adjustment. A third score is derived by summing the two independent scores for a total adjustment score. Thus, there were three scores for each student.

Reliability coefficients for the various parts of the personality tests were taken from the Manual (147, p. 4). The ranges are .80 to .88 for the primary level and .92 to .94 for the elementary level, for the three scores analyzed.

The independent variables considered in the study were sex, grade-level, and organization. Three hundred forty-nine students in grades 1-6, 179 boys, and 170 girls, took the personality tests in the first year. In the second year 381 students, 186 boys and 195 girls took the personality test. These distributions are shown in Tables 44 and 45.

TABLE 44

Distribution of Students Participating in Personality
Tests by Grade, Sex, and Organization

Academic Year 1961-1962

	Washington Team		Washington Self-Contained		Longfellow Self-Contained	
Grade	Male	Female	Male	Female	Male	Female
1	7	9	9	7	11	7
2	9	8	10	8	7	11
3	9	7	11	9	14	7
4	9	11	9	9	15	9
5	9	9	6	11	16	14
6	9	8	6	8	13	18
Sub Totals by Sex and Organization	52	52	51	52	76	66
Sub Totals by Organization	104		103		142	
Total N	349					

TABLE 45

Distribution of Students Participating in Personality Tests by Grade, Sex, and Organization

Academic Year 1962-1963

Grade	Washington Team Male	Washington Team Female	Washington Self-Contained Male	Washington Self-Contained Female	Longfellow Self-Contained Male	Longfellow Self-Contained Female
1	11	5	11	10	12	23
2	6	8	11	6	14	18
3	11	9	9	9	11	13
4	8	8	9	8	14	11
5	9	11	8	10	14	13
6	7	9	7	9	14	15
Sub Totals by Sex and Organization	52	50	55	52	79	93
Sub Totals by Organization	102		107		172	
Total N			381			

III. PERSONALITY: STATISTICAL ANALYSIS

A. Academic Year 1961-1962

Tables 46-48, below, show the pre- and post-test mean scores for personal, social, and total adjustment, by sex, grade, and organization from the first year personality tests. Raw scores for the primary levels were transformed to a scale comparable to the intermediate level scores, by means of a standard score transformation.

TABLE 46

California Test of Personality: Personal Adjustment Pre and Post Test Mean Scores by Sex, Grade, and Organization

Academic Year 1961-1962

		Washington Team		Washington Self-Contained		Longfellow Self-Contained	
Sex	Grade	Pre-Test	Post-Test	Pre-Test	Post-Test	Pre-Test	Post-Test
Girls	1	48.1	40.0	43.4	48.9	51.4	44.7
	2	54.0	53.9	45.5	49.1	44.6	47.0
	3	56.3	54.0	53.2	53.0	47.4	62.7
	4	44.7	46.7	53.3	52.7	55.6	56.9
	5	44.2	51.4	55.9	53.7	49.3	55.1
	6	46.4	52.1	48.8	58.5	47.7	53.4
Boys	1	50.3	48.0	46.8	56.4	48.9	53.4
	2	46.6	56.1	47.6	51.1	56.4	53.4
	3	47.0	43.8	49.0	55.5	45.4	56.4
	4	45.6	46.6	39.7	42.2	48.7	50.7
	5	57.8	53.6	41.3	46.7	46.4	52.3
	6	52.1	53.2	53.7	48.5	46.2	52.2

TABLE 47

California Test of Personality: Social Adjustment Pre and Post Test Mean Scores by Sex, Grade, and Organization

Academic Year 1961-1962

		Washington Team		Washington Self-Contained		Longfellow Self-Contained	
Sex	Grade	Pre-Test	Post-Test	Pre-Test	Post-Test	Pre-Test	Post-Test
Girls	1	50.6	44.9	52.6	60.1	62.4	51.7
	2	59.0	59.9	57.9	53.4	54.1	53.6
	3	60.0	63.4	54.7	55.6	60.1	65.1
	4	53.3	48.2	57.8	61.8	59.7	61.7
	5	54.8	56.9	61.7	61.0	56.4	60.6
	6	54.5	53.0	56.3	62.1	55.1	60.2
Boys	1	51.1	54.1	49.7	58.2	55.5	52.7
	2	49.9	52.9	55.7	50.4	57.3	55.1
	3	50.2	54.3	53.4	56.7	55.9	58.6
	4	51.9	48.4	47.6	48.0	54.3	50.5
	5	60.4	55.2	44.0	50.3	49.9	54.0
	6	58.4	54.0	55.0	50.5	53.7	52.8

TABLE 48

California Test of Personality: Total Adjustment Pre and Post Test Mean Scores by Sex, Grade, and Organization

Academic Year 1961-62

Sex	Grade	Washington Team		Washington Self-Contained		Longfellow Self-Contained	
		Pre-Test	Post-Test	Pre-Test	Post-Test	Pre-Test	Post-Test
Girls	1	98.7	84.9	96.0	109.0	113.8	96.4
	2	113.1	113.8	103.4	102.4	98.7	100.6
	3	116.4	117.4	107.9	108.6	107.5	127.8
	4	98.0	94.9	111.1	114.4	115.2	118.6
	5	99.0	108.3	117.6	114.7	105.7	115.7
	6	100.9	105.1	105.0	120.6	102.8	113.7
Boys	1	101.3	102.1	96.5	114.6	104.4	106.1
	2	96.4	109.0	103.3	101.5	113.6	113.5
	3	97.2	98.1	102.4	112.2	101.3	115.0
	4	97.4	95.0	87.2	90.2	102.9	101.3
	5	118.2	108.8	85.3	97.0	96.2	106.3
	6	110.6	107.2	108.7	99.0	99.9	105.1

Adjusted means were derived from the Bi-Md-14 Program which covaried out pre-test scores.* The F-ratios computed from the adjusted means are shown in Table 49. (Separate analyses were computed for the three scores, personal, social, and total adjustment.)

TABLE 49

F-Ratios of California Test of Personality Scores
by Type of Adjustment and Source

Academic Year 1961-62

	Type of Adjustment		
Source	Personal	Social	Total
Organization	4.38*	1.56	2.99
Grade	1.74	2.15	1.91
Sex	-	4.03*	-
Org x Grade	2.27*	2.25*	2.51*
Org x Sex	-	-	-
Grade x Sex	2.73*	1.96	3.28**
Org x Grade x Sex	-	-	1.10

*$p < .05$
**$p < .01$

---

* These adjusted means are given in Appendix B.

The significance of organization for personal adjustment scores was caused by the large difference between Longfellow Self-Contained and Washington Team for overall gain. These scores are shown in Table 50 below. The Duncan minimum range estimates show the difference which is required between organization means for significance.

TABLE 50

California Test of Personality: Personal Adjustment
Overall Mean Scores by Organization

Academic Year 1961-62

Scores	Washington Team	Washington Self-Contained	Longfellow Self-Contained	Duncan Range WT-WSC	Duncan Range WT-LSC	Duncan Range WSC-LSC
Pre Test	49.164	48.495	48.395			
Post Test	49.788	51.708	53.352			
Gain	.624	3.213	4.957	2.74	2.67	2.55

The significance of sex for social adjustment scores was caused by the fact that girls rose .605 in mean score between pre- and post-tests while boys rose only .196 in mean score. The raw score means in social adjustment were 56.507 and 57.112 for girls on pre- and post-tests respectively, 53.145 and 53.341 for boys.*

Sources of the interactions may be seen more clearly in the adjusted mean scores which have already taken into account variance due to pre-test. Figures 14-19, which follow, show graphs of these adjusted means. (Tables of the adjusted means for the California Personality Tests, 1961-62, are given in Appendix B).

---

* The authors are aware, however, that the statistical significance shown here is unmeaningful in terms of the actual test where this difference is less than a single question.

The sex by grade interaction for personal adjustment is shown in Figure 14: Girls increased in personal adjustment between grades 2 and 3, and 5 and 6, while boys decreased between these grades. Corresponding interactions do not show significance for social adjustment, shown in Figure 15, because of the larger variation in social adjustment scores. The significant interaction in total adjustment scores is probably mainly due to the differences between grades 5 and 6, (Figure 16).

Figure 17 shows the probable source of the grade by organization interaction for personal adjustment. While the Washington Team increased in personal adjustment between the first and second grades, decreased between the second and third grades, and increased between the third and fourth, Washington School Self-Contained did just the reverse: that is it decreased between the first and second grades, increased between the second and third, and decreased between the third and fourth. Longfellow School increased in personal development as did the Team between first and second grades, but followed the pattern of Washington School Self-Contained classes in the following two grades.

Figure 18 shows that the same interaction patterns occurred between organizations for social adjustment, between the first and second grades but disappeared on higher grade levels. Thus, total adjustment, shown in Figure 19,

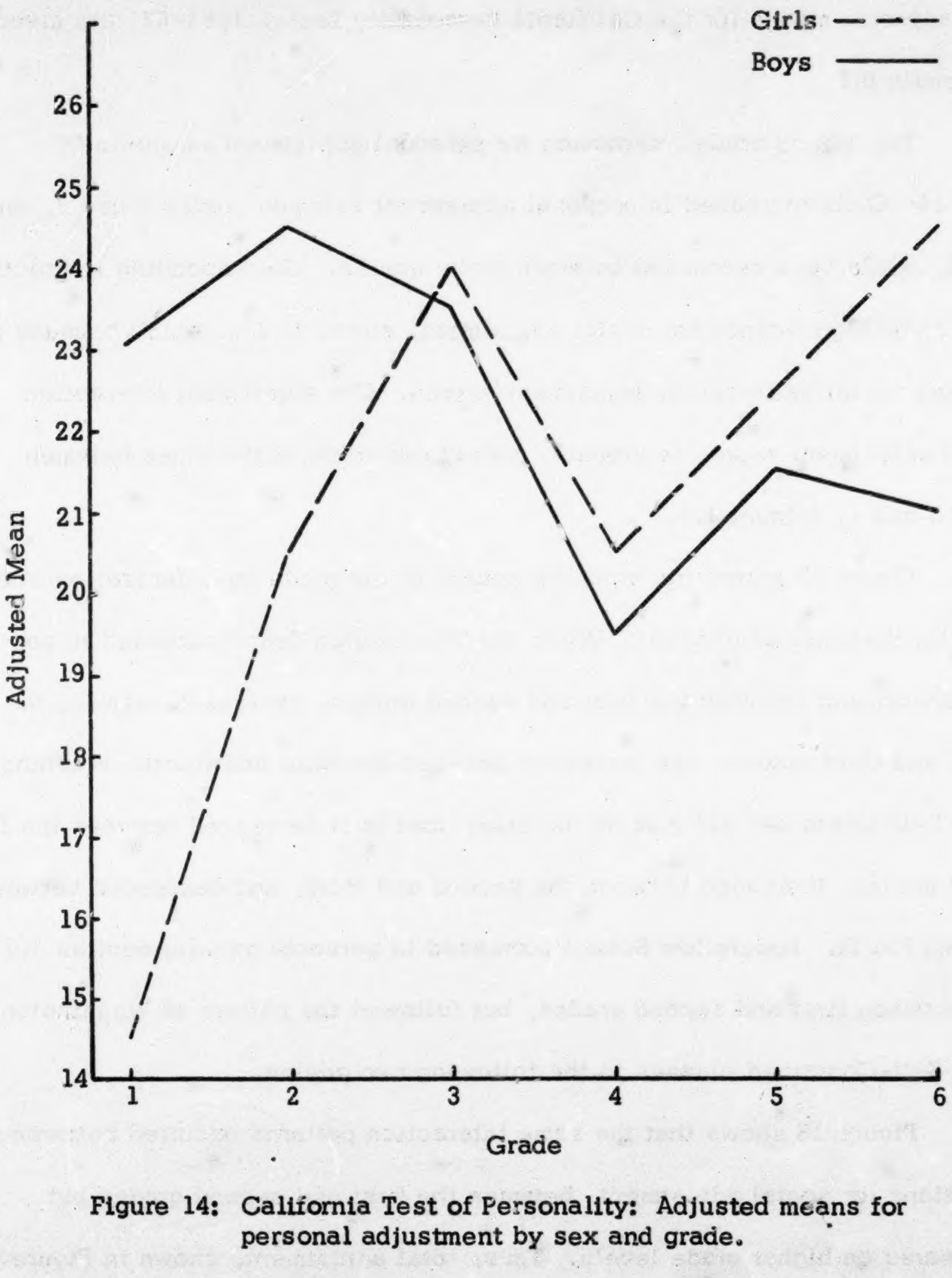

Figure 14: California Test of Personality: Adjusted means for personal adjustment by sex and grade.

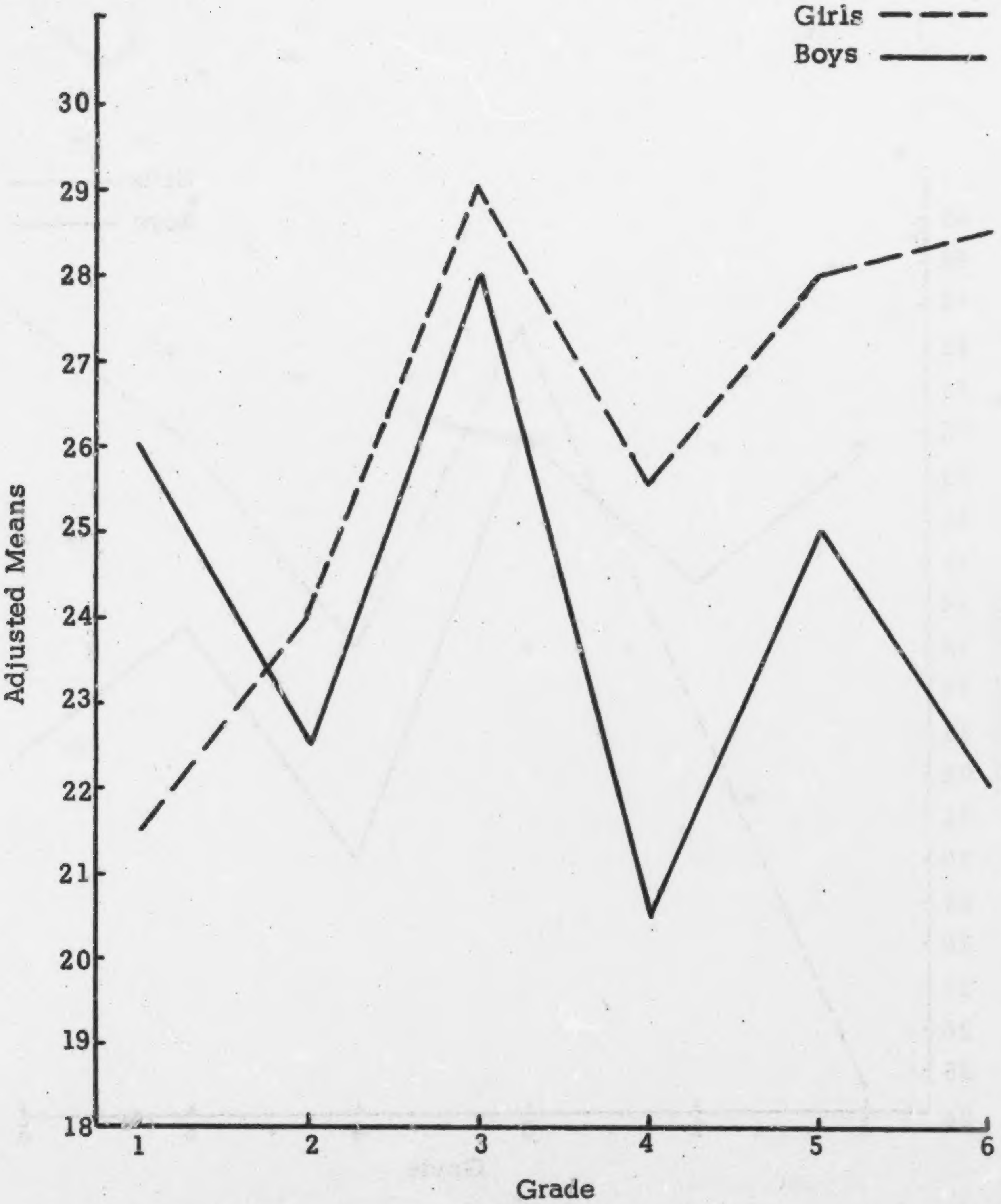

Figure 15: California Test of Personality: Adjusted means for social adjustment by sex and grade.

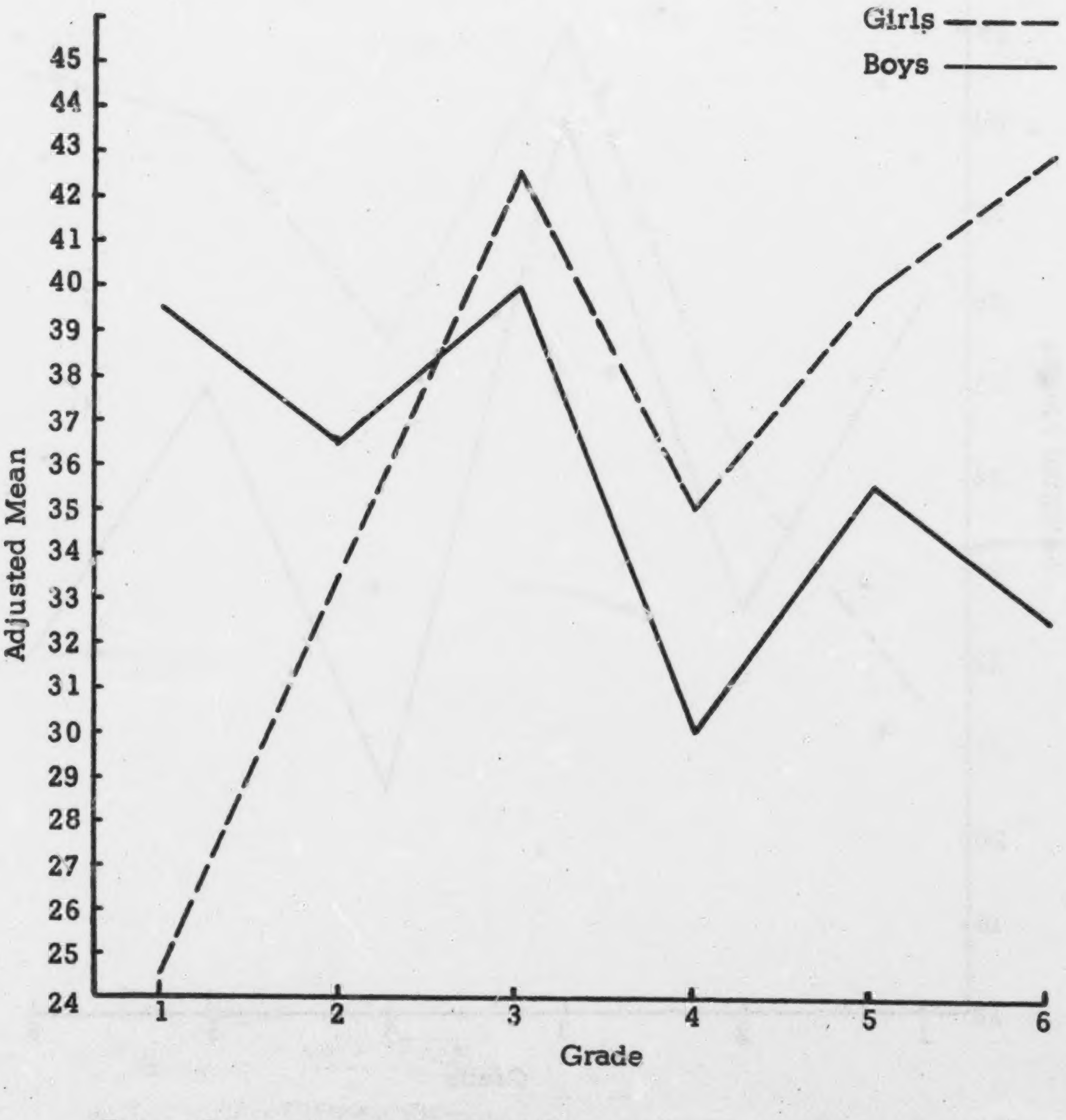

Figure 16: California Test of Personality: Adjusted means for total adjustment by sex and grade.

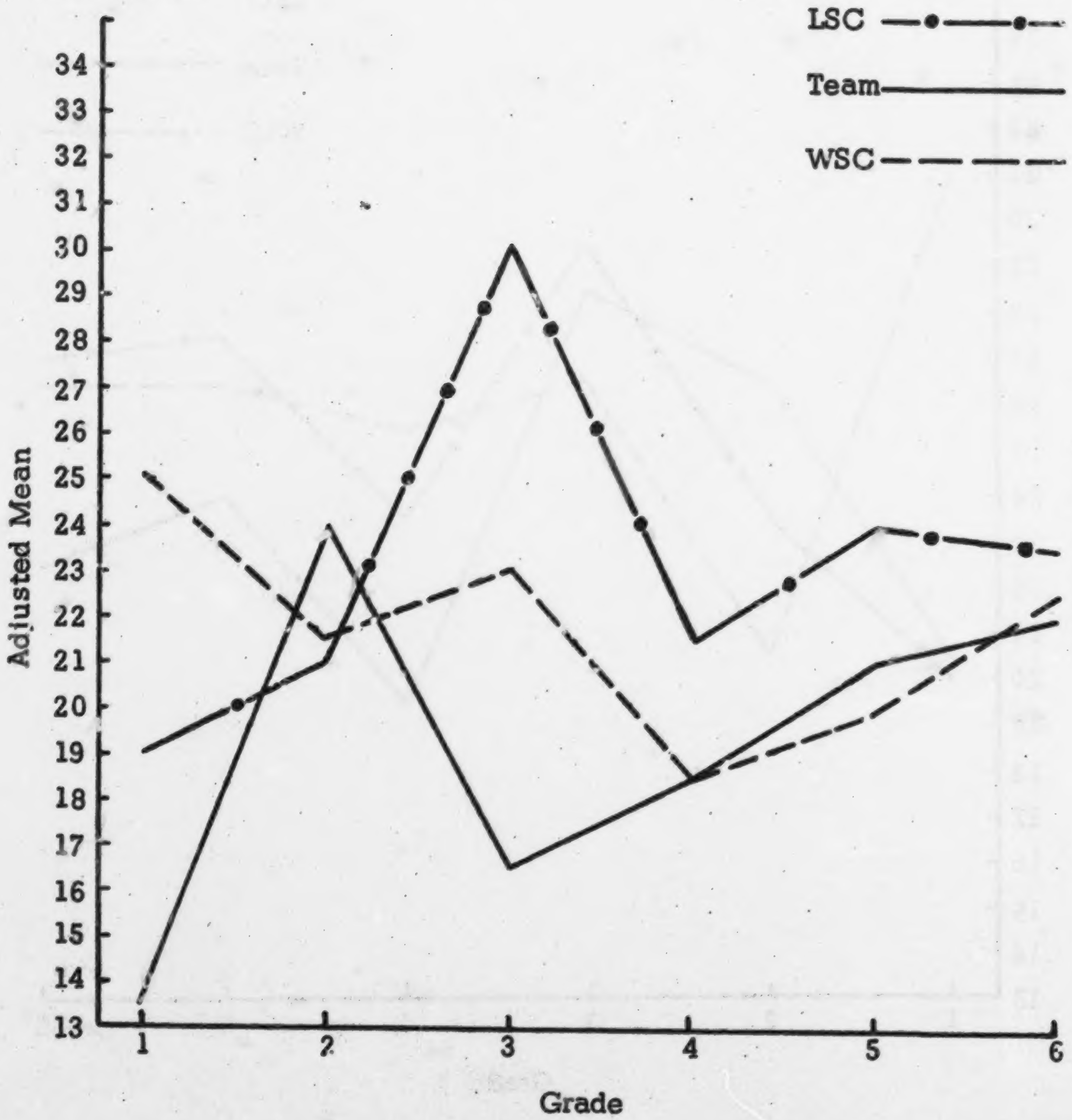

Figure 17: California Test of Personality: Adjusted means for personal adjustment by organization and grade.

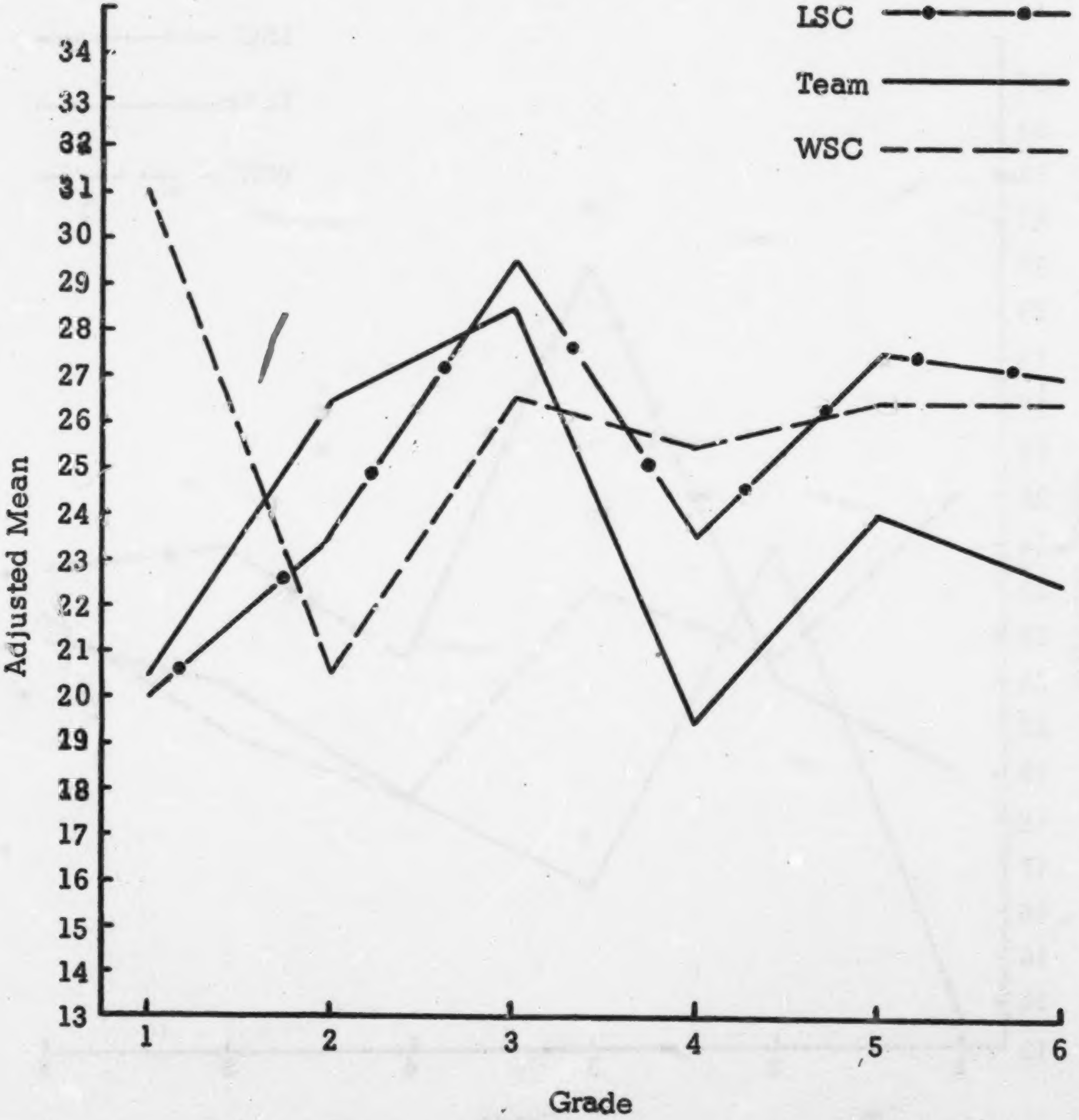

Figure 18: California Test of Personality: Adjusted means for social adjustment by organization and grade.

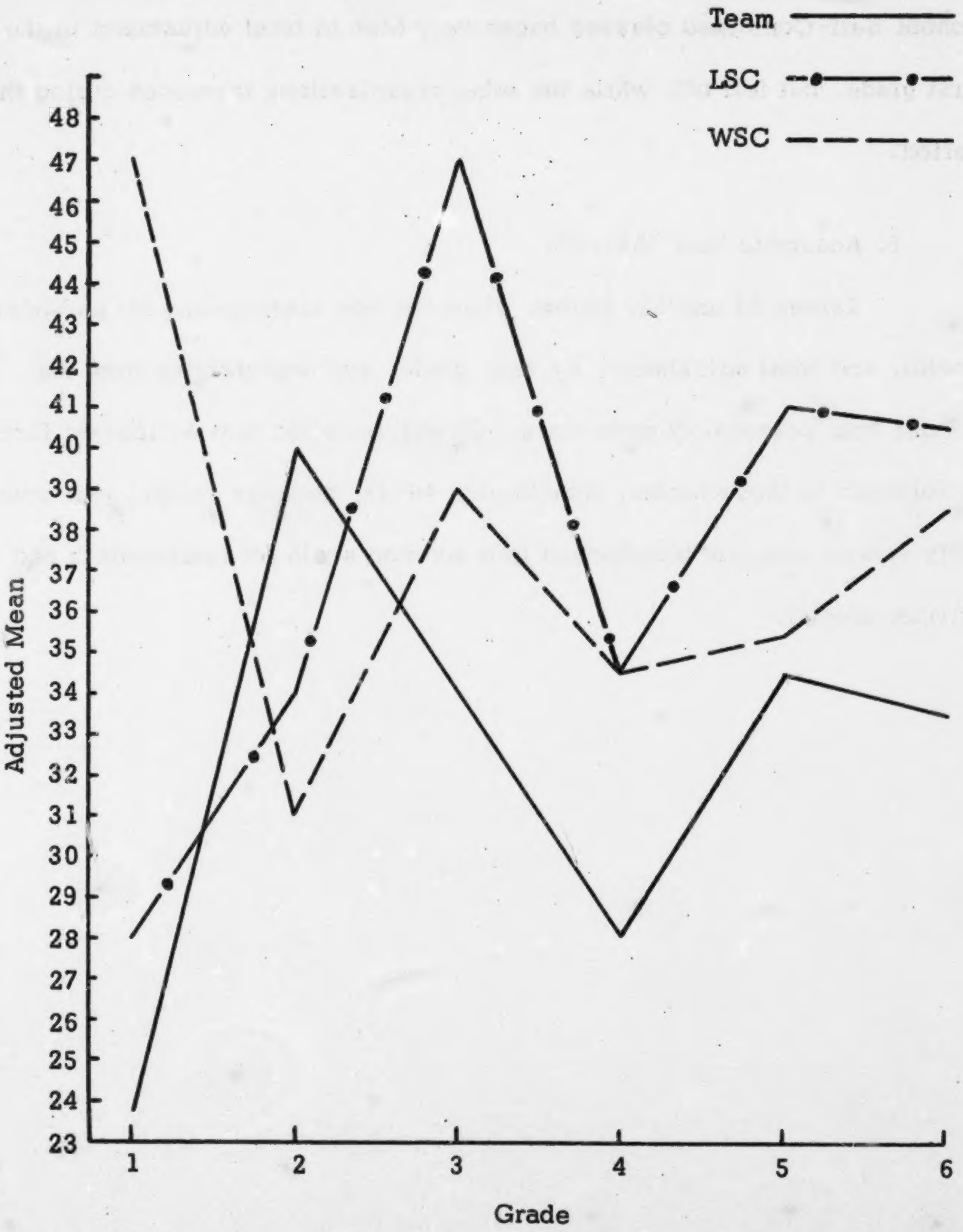

Figure 19: California Test of Personality: Adjusted means for total adjustment by organization and grade.

primarily reflects the interaction of the first and second grades: Washington School Self-Contained classes began very high in total adjustment in the first grade, but fell off, while the other organizations increased during this period.

B. Academic Year 1962-63.

Tables 51 and 52, below, show the raw score means for personal, social, and total adjustment, by sex, grade, and organization from the second year personality post-tests. (It will be noted that a different format is followed in these tables, from Tables 46-48, because second year personality scores were not transformed to a common scale for intermediate and primary levels).

TABLE 51

California Teat of Personality Mean Scores by Sex, Grade, and Organization; Primary Level

Academic Year 1962-1963

Area of Test	Sex	Grade	Washington Team	Washington Self-Contained	Longfellow Self-Contained
Personal Adjustment	Girls	1	31.600	31.100	33.174
		2	28.375	30.500	33.389
		3	32.889	31.778	32.077
	Boys	1	36.000	33.909	34.167
		2	34.000	31.545	34.357
		3	34.636	35.556	31.909
Social Adjustment	Girls	1	31.800	31.500	34.609
		2	32.375	36.333	36.667
		3	34.889	36.222	33.923
	Boys	1	34.091	31.000	36.250
		2	34.000	32.545	34.000
		3	36.455	34.111	33.273
Total Adjustment	Girls	1	63.400	62.600	67.783
		2	60.750	66.833	70.056
		3	67.778	68.000	66.000
	Boys	1	70.091	64.909	70.417
		2	68.000	64.090	68.357
		3	71.091	69.667	65.182

TABLE 52

California Test of Personality Mean Scores by Sex, Grade, and Organization; Intermediate Level

Academic Year 1962-1963

Area of Test	Sex	Grade	Washington Team	Washington Self-Contained	Longfellow Self-Contained
Personal Adjustment	Girls	4	52.500	48.375	56.455
		5	47.909	55.700	52.692
		6	51.111	56.889	55.133
	Boys	4	46.500	48.556	53.000
		5	55.778	54.250	48.286
		6	55.714	51.571	53.643
Social Adjustment	Girls	4	58.375	53.750	61.273
		5	54.818	59.800	58.231
		6	56.556	64.222	58.267
	Boys	4	53.250	52.000	58.214
		5	60.333	54.625	49.143
		6	55.571	51.571	55.786
Total Adjustment	Girls	4	110.875	102.125	117.727
		5	102.727	115.500	110.923
		6	107.667	121.111	113.400
	Boys	4	99.750	100.556	111.214
		5	116.111	108.875	97.429
		6	111.286	103.143	109.429

Analyses of variance were computed separately on the primary and intermediate levels for each of the three scores. The division of these analyses is therefore 2 x 3 x 3 (two sexes,, three organizations, three grades) as opposed to the 2 x 3 x 6 division of the first year covariance analyses. The resulting F-ratios are shown below in Tables 53 and 54.

TABLE 53

F-Ratios of California Test of Personality Scores on the Primary Level by Type of Adjustment and Source

Academic Year 1962-1963

	Type of Adjustment		
Source	Personal	Social	Total
Organization	-	-	-
Grade	-	-	-
Sex	4.28*	-	1.07
Org x Grade	-	1.21	-
Org x Sex	-	1.28	-
Grade x Sex	-	-	-
Org x Grade x Sex	-	-	-

*p<.05

TABLE 54

F-Ratios of California Test of Personality Scores on the Intermediate Level by Type of Adjustment and Source

Academic Year 1962-1963

	Type of Adjustment		
Source	Personal	Social	Total
Organization	–	–	–
Grade	1.09	–	–
Sex	–	7.28**	2.72
Org x Grade	1.11	1.82	1.52
Org x Sex	–	1.75	1.34
Grade x Sex	–	–	–
Org x Grade x Sex	–	1.84	1.44

**$p < .01$

The significant main effects were sex, for personal adjustment on the primary level, and for social adjustment on the intermediate level. Girls were lower than boys in personal adjustment on the primary level, with respective means of 32.099 and 33.989, while they were higher than boys in social adjustment on the intermediate level, with respective means of 58.405 and 54.533.

IV. ACHIEVEMENT: MEASUREMENT

The California Achievement Tests (1957 edition) by E. W. Tiegs and W. W. Clark, were used for the 1961-62 study. Four test levels were used with two forms on each level, one for pre-tests, and the other for post-tests. The test level corresponding to each grade is shown in Table 55.

TABLE 55

Achievement Test by Grade, Test Level and Date

Grade	October 1961 Level Taken as Pre-Test	May 1962 Level Taken as Post-Test
1	Lower Primary Form W	Lower Primary Form X
2	Upper Primary Form W	Upper Primary Form X
3	Upper Primary Form W	Upper Primary Form X
4	Elementary Form W	Elementary Form X
5	Elementary Form W	Elementary Form X
6	Junior High Form W	Junior High Form X

In the second year of study, the California Achievement Tests were again used for the first and second grades, while the Iowa Test of Basic Skills was used in the remaining grades. Tests and test levels corresponding to each grade are shown in Table 56.

California Achievement Tests supply scores for arithmetic, reading, language, and a total score. The Iowa Test of Basic Skills, however, gives fifteen scores, some of which are combinations of single scores. Only six of these scores, however, were used in the analysis: vocabulary; reading; total language; total work-study skills; total arithmetic; and, composite. The six scores cover all areas of the test.

TABLE 56

Achievement Test by Grade, Test Level and Date

May, 1963

Grade	Test	Level
1	California Achievement	Lower Primary Form W
2	California Achievement	Upper Primary Form W
3	Iowa Basic Skills	Form 1
4	Iowa Basic Skills	Form 1
5	Iowa Basic Skills	Form 1
6	Iowa Basic Skills	Form 2*

* Form 2 was used because the sixth grade students had been tested on Form 1 six months previously.

The reliability coefficients for the California Test were found in the manuals. Reliability coefficients are given for each part of each level. The range for each level is reported in Table 57.

TABLE 57

Range of Reliability Coefficients on the Different Levels of the California Achievement Tests

Level	Reliability Coefficients Low	High
Lower Primary	.78	.93
Upper Primary	.69	.95
Elementary	.70	.95
Junior High	.83	.95

The extraordinary length of the Iowa Test of Basic Skills increases administration problems but at the same time increases the reliability of the measures. The authors of the Iowa Test of Basic Skills report that reliability "has been a first consideration" in constructing the tests. * The test authors state that any attempt to produce a short-form of the Iowa Test would produce results unreliable for individual pupil diagnosis.

The reliability coefficients reported were computed by the split-halves procedure, each coefficient being based on five hundred cases drawn randomly from the complete standardization population at that grade level. The reliability coefficients are substantial in all areas as shown in the table below.

---

* Manual for Administrators, Supervisors, and Counselors; Iowa Test of Basic Skills; Houghton Mifflin Company, Boston, 1956, p. 12.

TABLE 58

Iowa Test of Basic Skills Reliability Coefficients by Subject Area and Grade

Test	Grade 3	 4	 5	 6
Vocabulary	.84	.86	.92	.92
Reading	.91	.96	.94	.93
Total Language	.93	.95	.95	.95
Total Work-Study Skills	.90	.93	.94	.94
Total Arithmetic	.89	.90	.89	.90
Composite	.97	.98	.98	.98

In the Achievement test analyses, sex, grade level, and organization were treated as independent variables. Distributions of students in these categories for each year of study are shown in Tables 59 and 60. (Because Grade 1 was deleted from the first year analysis, it will be noted that there are 30 cells in the 1961-62 Distribution, Table 59, while there are 36 cells in the 1962-63 Distribution, Table 60.)

In the two schools, 152 boys and 147 girls, grades 2-6, participated in the first year study, and 186 boys and 175 girls, grades 1-6, participated in the second year study.

TABLE 59

Distribution of Students Participating in Achievement Tests by Grade, Sex, and Organization

Academic Year 1961-62

	Washington Team		Washington Self-Contained		Longfellow Self-Contained	
Grade	Male	Female	Male	Female	Male	Female
2	9	8	10	8	7	11
3	9	7	11	9	14	7
4	9	11	9	9	15	9
5	9	9	6	11	16	14
6	9	8	6	8	13	18
Subtotals by Sex and Organization	45	43	42	45	65	59
Subtotals by Organization	88		87		124	
Total N			299			

TABLE 60

Distribution of Students Participating in Achievement Tests by Grade, Sex, and Organization

Academic Year 1962-63

	Washington Team		Washington Self-Contained		Longfellow Self-Contained	
Grade	Male	Female	Male	Female	Male	Female
1	11	5	11	10	12	23
2	6	8	11	6	14	18
3	11	9	9	9	11	13
4	8	8	9	8	14	11
5	9	10	8	10	14	13
6	7	10	7	9	14	15
Subtotals by Sex and Organization	52	50	55	52	79	93
Subtotals by Organization	102		107		172	
Total N			381			

V. ACHIEVEMENT: STATISTICAL ANALYSIS

A. Academic Year 1961-62

Tables 61-64, below, show the pre- and post-test scores for reading, arithmetic, language, and total achievement, by sex, grade, and organization from the first year California Achievement Tests.

TABLE 61

California Achievement Test: Reading Mean Grade Placement, Pre- and Post-Tests by Sex, Grade, and Organization: Grades 2-6

Academic Year 1961-62

		Washington Team		Washington Self-Contained		Longfellow Self-Contained	
Sex	Grade	Pre-Test	Post-Test	Pre-Test	Post-Test	Pre-Test	Post-Test
Girls	2	2.59	3.71	2.88	3.63	2.77	3.73
	3	4.09	4.70	4.21	4.96	3.99	5.49
	4	5.95	6.18	5.46	6.01	6.31	6.62
	5	6.52	6.96	6.94	7.64	6.73	6.86
	6	7.37	7.34	6.90	7.69	8.17	8.53
Boys	2	2.36	2.95	2.33	2.89	2.36	3.59
	3	3.19	4.15	3.59	4.42	3.63	5.00
	4	5.12	5.55	5.09	5.64	6.14	6.19
	5	6.67	6.94	6.01	6.91	6.93	7.58
	6	7.18	6.96	7.64	7.87	7.85	7.95

TABLE 62

California Achievement Test: Arithmetic Mean Grade Placement, Pre- and Post-Tests by Sex, Grade, and Organization: Grades 2-6

Academic Year 1961-62

		Washington Team		Washington Self-Contained		Longfellow Self-Contained	
Sex	Grade	Pre-Test	Post-Test	Pre-Test	Post-Test	Pre-Test	Post-Test
	2	2.38	3.39	2.57	3.36	2.83	3.36
	3	3.25	4.14	3.38	4.25	3.61	5.25
Girls	4	4.86	5.90	4.87	6.09	5.06	5.63
	5	6.01	6.48	5.73	7.30	5.75	6.77
	6	6.24	6.58	7.10	9.10	7.31	9.12
	2	2.09	2.95	2.42	3.20	3.01	3.56
	3	2.91	3.89	3.06	4.20	3.46	5.27
Boys	4	4.36	4.86	4.80	5.48	5.54	5.97
	5	5.90	6.92	5.61	7.36	6.23	7.05
	6	6.89	6.91	7.13	9.50	7.09	8.87

TABLE 63

California Achievement Test: Language Mean Grade Placement, Pre- and Post-Tests by Sex, Grade, and Organization: Grades 2-6

Academic Year 1961-62

		Washington Team		Washington Self-Contained		Longfellow Self-Contained	
Sex	Grade	Pre-Test	Post-Test	Pre-Test	Post-Test	Pre-Test	Post-Test
Girls	2	2.77	3.91	2.86	3.84	2.91	3.82
	3	4.01	4.74	3.60	4.71	3.89	5.53
	4	6.19	6.58	5.42	6.19	6.69	6.83
	5	6.34	7.43	6.77	7.87	6.66	7.16
	6	6.78	7.14	7.03	8.18	7.89	7.96
Boys	2	2.07	3.23	2.31	3.04	2.84	3.81
	3	2.85	4.14	3.17	4.07	3.13	5.01
	4	4.80	5.08	4.67	5.46	5.85	6.20
	5	5.91	6.47	5.68	7.46	6.38	7.00
	6	6.30	6.36	6.51	6.93	7.14	7.32

TABLE 64

California Achievement Test: Total Achievement Mean Grade Placement, Pre- and Post-Tests by Sex, Grade, and Organization: Grades 2-6

Academic Year 1961-62

		Washington Team		Washington Self-Contained		Longfellow Self-Contained	
Sex	Grade	Pre-Test	Post-Test	Pre-Test	Post-Test	Pre-Test	Post-Test
Girls	2	2.58	3.67	2.77	3.61	2.84	3.64
	3	3.79	4.52	3.73	4.64	3.83	5.42
	4	5.67	6.21	5.25	6.10	6.02	6.36
	5	6.29	7.02	6.48	7.60	6.38	6.93
	6	6.80	7.02	7.01	8.33	7.79	8.53
Boys	2	2.17	3.04	2.35	3.04	2.73	3.65
	3	2.98	4.06	3.28	4.23	3.41	5.09
	4	4.76	5.16	4.78	5.53	5.85	6.12
	5	6.16	6.78	5.77	7.24	6.52	7.21
	6	6.79	6.74	7.09	8.10	7.36	8.05

An Analysis of Covariance was computed on each of the four scores, using the Bi-Md-14 Program. I.Q. and pre-test scores were treated as covariates. The division of this analysis is comparable to the analysis of the first year personality tests except that first grade was omitted from the study, as explained above.* Table 65 shows the resultant F-ratios.

---

* See page 115.

TABLE 65

F-Ratios of California Achievement Test Scores by Subject Area and Source

Academic Year 1961-62

	Subject Area			
Source	Reading	Arithmetic	Language	Total Achievement
Organization	5.69**	27.35***	3.84*	15.94***
Grade	3.82**	12.36***	7.83***	14.02***
Sex	1.73	-	-	-
Org x Grade	3.24**	17.24***	4.53***	12.97***
Org x Sex	-	-	1.47	-
Grade x Sex	1.21	2.63*	-	1.98
Org x Grade x Sex	-	1.39	1.31	-

$*p < .05$
$**p < .01$
$***p < .001$

The sources of organizational significance may be seen in Table 66 below which shows overall organizational pre- and post-test means, and total gains, for each score. The Duncan Range was actually computed on the gain scores using the results of the analysis. This is an approximation procedure although a reasonable one, since I.Q. turned out to be a randomly distributed covariate. This fact is illustrated at a later point in the report. The Duncan Range has been computed in each case to determine which differences are significant. (Use of the Duncan Range on gain scores, for the covariance analysis, is discussed above, page 114.)

TABLE 66

California Achievement Test: Overall Raw Mean Grade Placement by Organization, for Reading, Arithmetic, Language, and Total Scores

Academic Year 1961-1962

Scores	Tests	Washington Team	Washington Self-Contained	Longfellow Self-Contained	Duncan Ranges		
					WT-WSC	WT-LSC	WSC-LSC
Reading	Post-Test	5.58	5.64	6.46			
	Pre-Test	5.15	4.97	5.88			
	Gain	.43	.67	.58	.20	.17	.17
Arithmetic	Post-Test	5.25	5.77	6.45			
	Pre-Test	4.53	4.51	5.31			
	Gain	.72	1.26	1.14	.18	.15	.15
Language	Post-Test	5.55	5.65	6.31			
	Pre-Test	4.85	4.69	5.65			
	Gain	.70	.96	.66	.22	.20	.21
Total Achievement	Post-Test	5.46	5.68	6.41			
	Pre-Test	4.84	4.71	5.61			
	Gain	.62	.97	.80	.14	.12	.12

In all scores, Washington Self-Contained gained significantly more than Washington Team, and in language and total scores, Washington Self-Contained gained significantly more than Longfellow Self-Contained. In arithmetic and total score, Longfellow Self-Contained gained significantly more than the teams.

The cause of grade level significance is simply the fact that each grade achieved correspondingly higher than the previous grade.

Grade by organization interactions can be seen in the graphs of the adjusted cell means, Figures 20-23.* In all four scores the major crossover occurred between third and fourth grades: Longfellow Self-Contained classes were much higher than the classes of both other organizations in the third grade, for all scores, and dropped below both other organizations in the fourth grade for all scores.

Between first and second grades there were less marked crossovers. Washington Team was higher than the other organizations in second grade arithmetic (Figure 21), language (Figure 22), and total achievement (Figure 23) and dropped behind the other organizations in the subsequent grades.

Crossovers also occurred to a lesser extent between fifth and sixth grades, where Washington Team fell sharply in total score, while the other organizations increased. The crossover in arithmetic (Figure 21) is by far the greatest, and is probably a source of the high interaction effect significance in this subject.

It should be noted that the graph of the adjusted means for total achievement (Figure 23) is very similar to the net gain graph of the same score (Figure 24). This illustrates that I.Q. was a randomly distributed covariate.

---

* These means, adjusted for the covariates, pre-test score and I.Q., may be found in Appendix B.

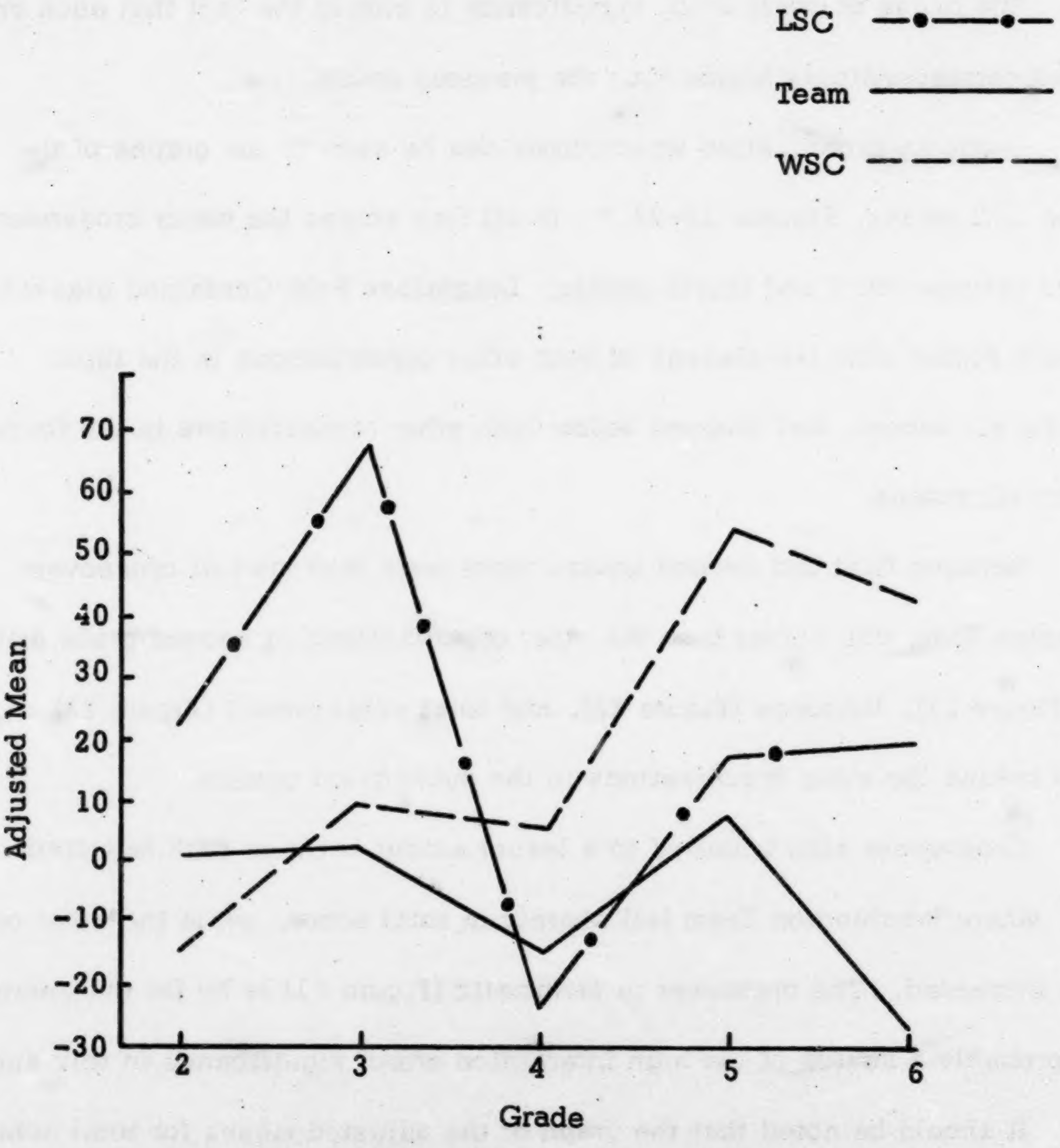

Figure 20: **California Achievement Test: Adjusted means for reading by grade and organization.**

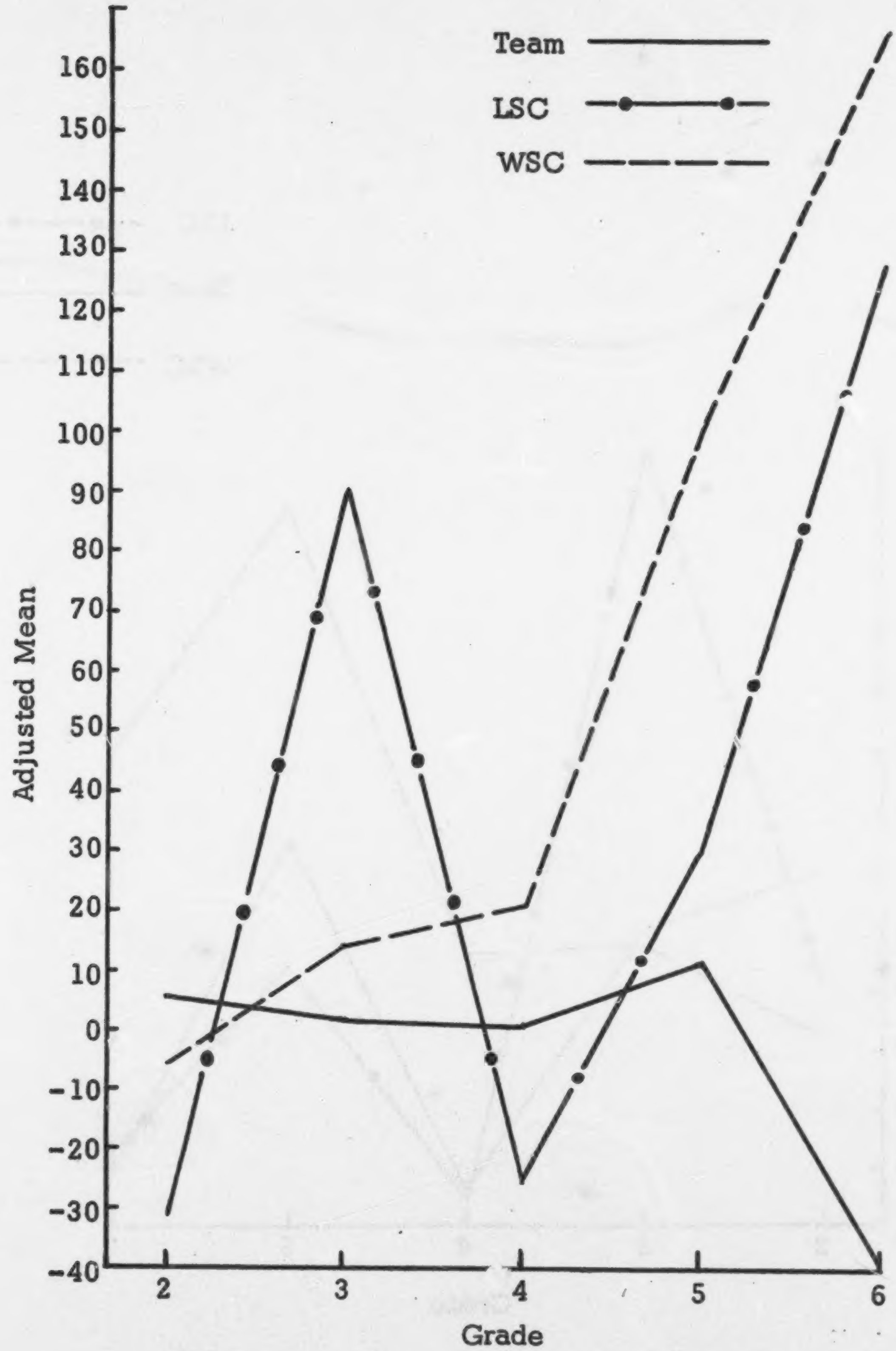

Figure 21: California Achievement Test: Adjusted means for arithmetic by grade and organization.

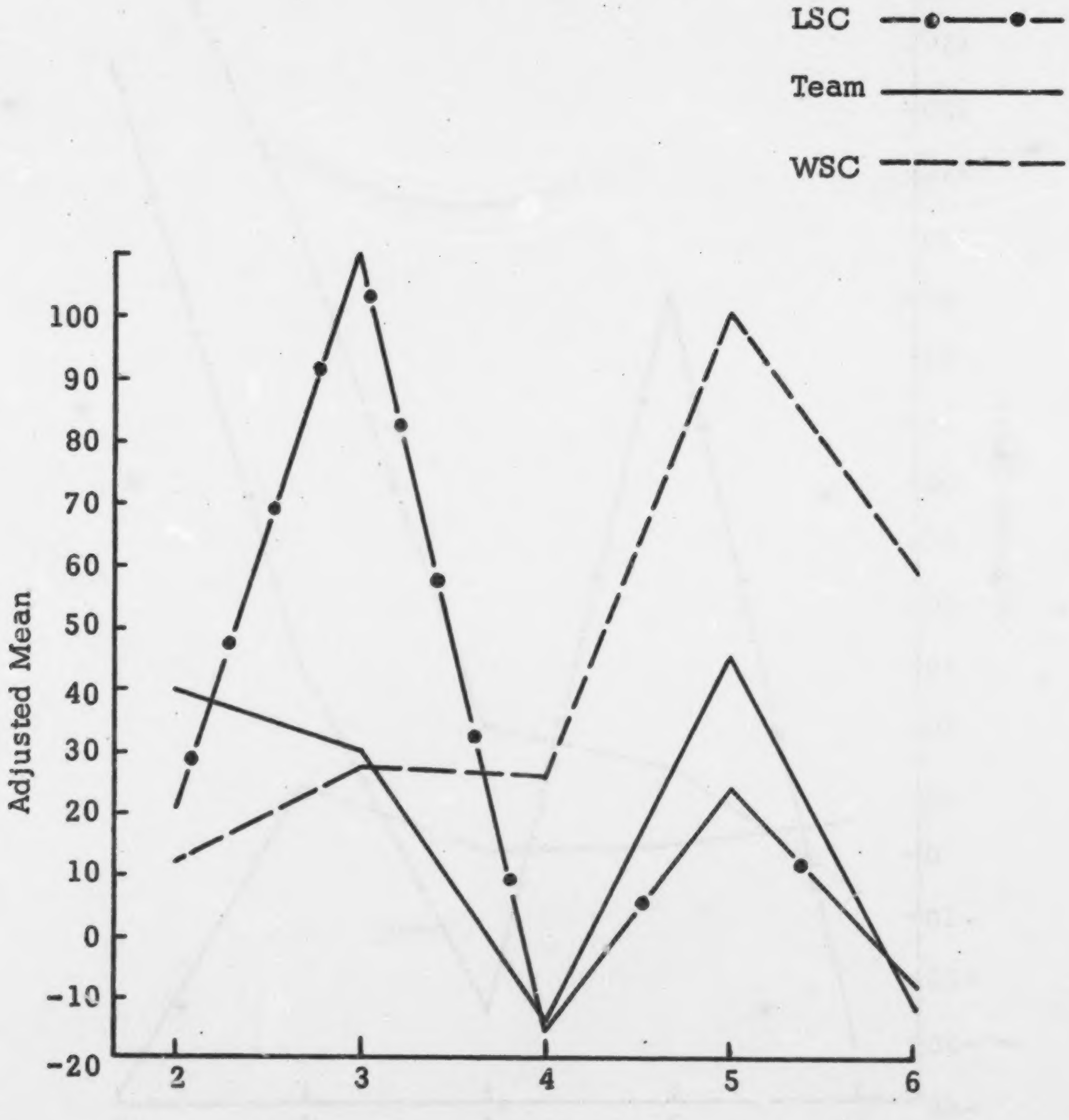

Figure 22: California Achievement Test: Adjusted means for language by grade and organization.

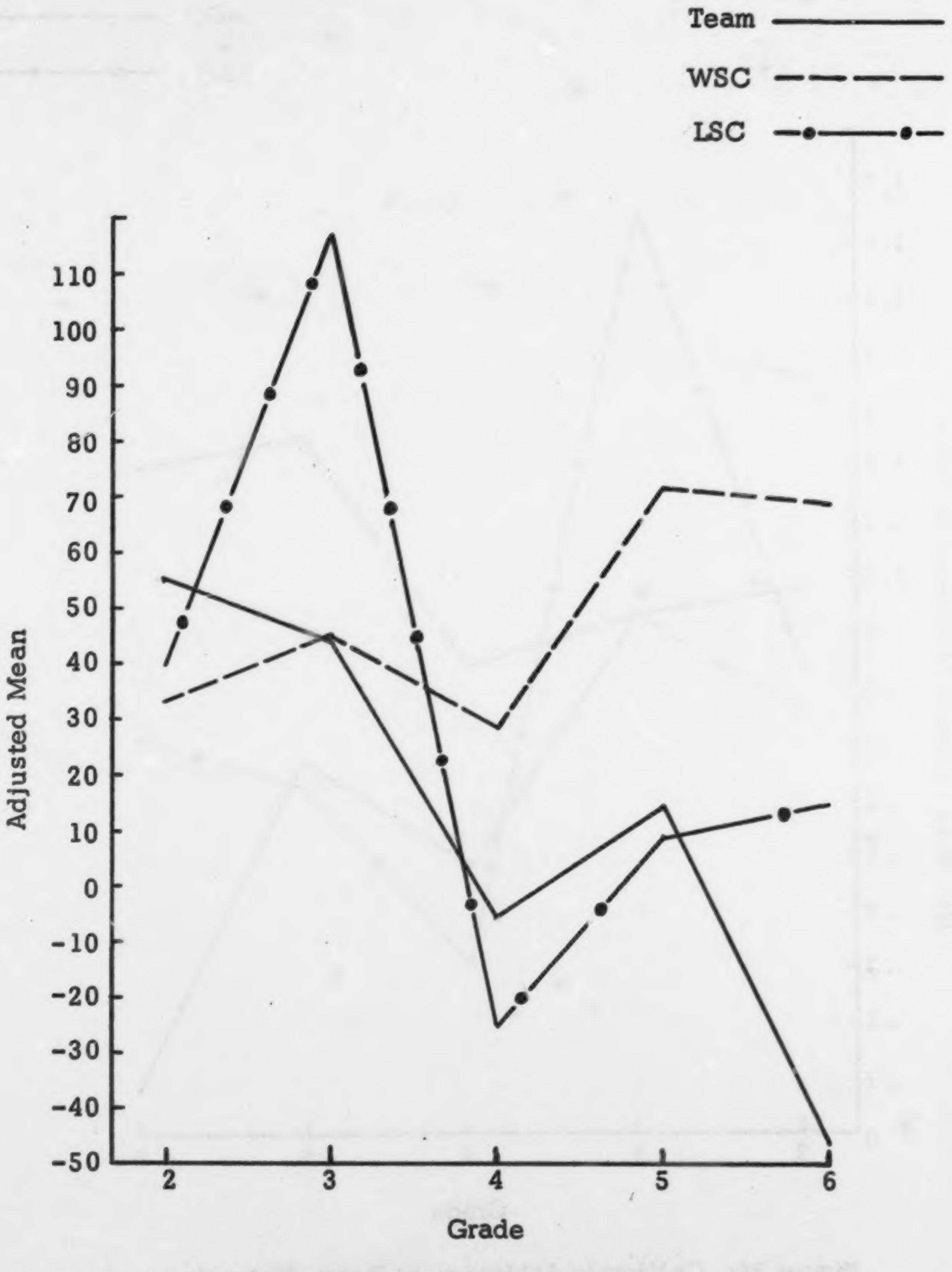

Figure 23: California Achievement Test: Adjusted means for total achievement by grade and organization.

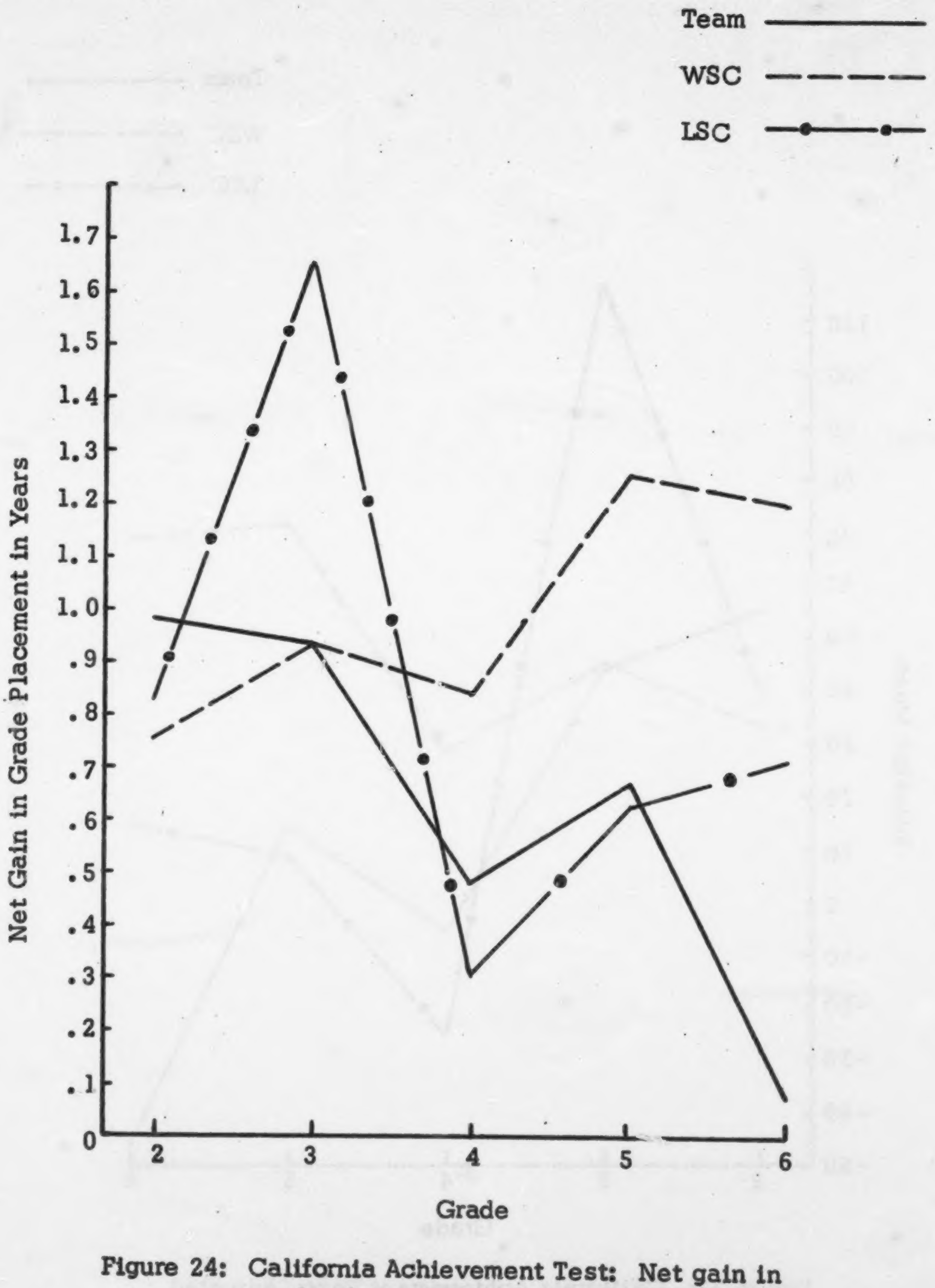

Figure 24: California Achievement Test: Net gain in grade placement for total achievement by grade and organization.

The interaction of grade and sex for arithmetic achievement is shown by the adjusted cell means given below in Table 67.

TABLE 67

California Achievement Test: Arithmetic
Adjusted Means by Sex and Grade

Academic Year 1961-62

Grade	Sex Girls	Boys
2	-13.488	- 8.826
3	25.202	47.971
4	20.520	-24.348
5	44.083	42.366
6	99.664	80.628

In the second and third grades, boys achieved better than girls, while they fell behind girls in the other three grades.

B. Academic Year 1962-63

Separate tests were given to the first grade, second grade, and third through sixth grades in the second year achievement study, as shown in Table 56, above. For the sake of clarity, these tests are treated separately in the present section.

1) First Grade Achievement

Table 68, below, shows the mean grade placements for reading, arithmetic, language, and total for the first grade on the California Achievement Tests.

TABLE 68

California Achievement Test: Reading, Arithmetic, Language and Total Achievement Mean Grade Placements by Sex, Grade, and Organization; First Grade

Academic Year 1962-1963

		Subject Area			
Sex	Organization	Reading	Arithmetic	Language	Total Achievement
	Washington Team	2.72	2.97	2.80	2.83
Girls	Washington Self-Contained	1.92	2.29	1.73	1.98
	Longfellow Self-Contained	2.57	2.50	2.87	2.64
	Washington Team	2.85	3.00	3.26	3.04
Boys	Washington Self-Contained	2.41	2.71	2.35	2.49
	Longfellow Self-Contained	3.00	2.79	3.29	3.03

Separate analyses of variance were computed for each subject area. The division of the first grade analyses was 2x3 for two sexes and three organizations, and the resultant F-ratios are shown below in Table 69.

TABLE 69

F-Ratios of California Achievement Test Scores for First Grade by Subject Area and Source

Academic Year 1962-63

Source	Subject Area Reading	Arithmetic	Language	Total Achievement
Organization	4.93*	2.75	10.66**	6.24*
Sex	3.57	2.02	5.93*	4.40*
Org x Sex	-	-	--	--

*p<.05
**p<.01

The sources of organizational significance may be seen in Table 70, below which shows the mean grade placements for reading, language, and total by organization.

TABLE 70

California Achievement Test: Reading, Language, and Total Achievement Overall Mean Grade Placements by Organization; First Grade

Academic Year 1962-63

Subject	Washington Team	Washington Self-Contained	Longfellow Self-Contained	Duncan Range WT-WSC	WT-LSC	WSC-LSC
Reading	2.81	2.18	2.71	.50	.43	.40
Language	3.12	2.06	3.01	.55	.48	.44
Total Achievement	2.97	2.25	2.77	.47	.41	.38

Washington Team was significantly higher than Washington Self-Contained in all scores, and Longfellow Self-Contained was significantly higher than Washington Self-Contained for all scores. Differences between Washington Team and Longfellow Self-Contained were non-significant.

Sex was significant in language and total achievement, as can be seen in the mean grade placements by Sex for these scores in Table 71.

Girls were lower than boys in language and total achievement.

There were no interaction effects in the first grade analysis of achievement.

TABLE 71

California Achievement Test: Language and Total Achievement Overall Mean Grade Placements by Sex; First Grade

Academic Year 1962-1963

	Sex	
Subject	Girls	Boys
Language	2.56	2.98
Total Achievement	2.49	2.86

2) Second Grade Achievement

Table 72, below, shows the mean grade placements for reading, arithmetic, language, and total for the second grade on the California Achievement Tests.

TABLE 72

California Achievement Test: Reading, Arithmetic, Language, and Total Achievement Mean Grade Placements by Sex, Grade, and Organization; Second Grade

Academic Year 1962-1963

		Subject Area			
Sex	Organization	Reading	Arithmetic	Language	Total Achievement
Girls	Washington Team	4.03	3.32	4.10	3.82
	Washington Self-Contained	3.27	3.08	2.99	3.12
	Longfellow Self-Contained	4.64	3.61	4.28	4.17
Boys	Washington Team	4.42	4.14	4.16	4.24
	Washington Self-Contained	3.57	3.29	3.44	3.43
	Longfellow Self-Contained	3.85	3.38	3.87	3.70

As in the first grade analysis, a 2x3 factorial ANOVA was computed for each subject. These F-ratios are shown in Table 73.

TABLE 73

F-Ratios of California Achievement Test Scores for Second Grade by Subject Area and Source

Academic Year 1962-63

	Subject Area			
Source	Reading	Arithmetic	Language	Total Achievement
Organization	3.65*	2.86	5.31**	4.14*
Sex	-	2.03	-	-
Org x Sex	1.75	2.65	-	1.44

$^{*}p < .05$

$^{**}p < .05$

The sources of organizational significance for reading, language, and total achievement scores can be seen in the table of overall mean grade placements, Table 74, below.

In these scores, as in the first grade scores, Longfellow Self-Contained is always significantly higher than Washington Self-Contained and never significantly different from Washington Team. However, Washington Team is only significantly higher than Washington Self-Contained in one score for the second grade, language.

There were no significant main effects for sex, and no significant interactions in this analysis.

TABLE 74

California Achievement Test: Reading, Language, and Total Achievement Overall Mean Grade Placements by Organization; Second Grade

Academic Year 1962-63

Subject	Washington Team	Washington Self-Contained	Longfellow Self-Contained	Duncan Range		
				WT-WSC	WT-LSC	WSC-LSC
Reading	4.20	3.46	4.30	.80	.75	.70
Language	4.13	3.28	4.10	.76	.68	.61
Total Achievement	4.00	3.32	3.97	.69	.62	.55

3) Third through Sixth Grade Achievement

Tables 75-80, below, show the mean grade placements for vocabulary, reading, total language, total work study skills, total arithmetic, and composite for the Iowa Test of Basic Skills administered to grades 3-6, in May, 1963.

TABLE 75

Iowa Test of Basic Skills: Vocabulary Mean Grade Placement by Sex, Grade and Organization Third-Sixth Grades

Academic Year 1962-63

Sex	Grade	Washington Team	Washington Self-Contained	Longfellow Self-Contained
Girls	3	4.33	4.38	4.68
	4	5.81	4.50	5.31
	5	5.95	5.62	6.92
	6	7.64	8.04	7.89
Boys	3	3.75	3.63	4.45
	4	4.39	4.78	5.44
	5	6.11	5.36	6.84
	6	8.07	6.73	8.66

TABLE 76

Iowa Test of Basic Skills: Reading Mean Grade Placement by Sex, Grade and Organization Third-Sixth Grades

Academic Year 1962-63

Sex	Grade	Washington Team	Washington Self-Contained	Longfellow Self-Contained
Girls	3	4.03	4.33	4.87
	4	5.66	4.79	4.86
	5	5.78	6.06	6.99
	6	7.42	7.21	7.11
Boys	3	3.64	3.52	4.19
	4	4.59	4.60	5.02
	5	6.21	5.38	6.42
	6	7.43	6.39	8.45

TABLE 77

Iowa Test of Basic Skills: Total Language Mean Grade Placement by Sex, Grade and Organization Third-Sixth Grades

Academic Year 1962-63

Sex	Grade	Washington Team	Washington Self-Contained	Longfellow Self-Contained
Girls	3	4.22	4.02	4.65
	4	5.91	4.78	5.49
	5	6.09	6.78	7.51
	6	7.63	8.39	8.11
Boys	3	3.53	3.27	4.31
	4	4.14	4.39	5.49
	5	5.71	5.53	6.50
	6	6.76	6.96	7.92

TABLE 78

Iowa Test of Basic Skills: Total Work Study Skills Mean Grade Placement by Sex, Grade and Organization Third-Sixth Grades

Academic Year 1962-63

Sex	Grade	Washington Team	Washington Self-Contained	Longfellow Self-Contained
Girls	3	3.91	4.02	4.65
	4	5.00	4.76	4.78
	5	5.70	6.02	6.36
	6	7.27	7.79	7.43
Boys	3	3.79	3.52	4.30
	4	4.14	4.39	5.09
	5	5.68	5.35	6.01
	6	7.11	6.94	7.81

TABLE 79

Iowa Test of Basic Skills: Total Arithmetic Mean Grade Placement by Sex, Grade, and Organization Third-Sixth Grades

Academic Year 1962-63

Sex	Grade	Washington Team	Washington Self-Contained	Longfellow Self-Contained
Girls	3	3.87	3.74	4.35
	4	4.88	4.45	5.05
	5	5.40	6.02	6.38
	6	6.83	7.72	7.56
Boys	3	3.45	3.49	4.28
	4	4.23	4.51	4.86
	5	5.47	5.45	5.86
	6	6.76	6.97	7.99

TABLE 80

Iowa Test of Basic Skills: Composite Mean Grade Placement by Sex, Grade, and Organization Third-Sixth Grades

Academic Year 1962-63

Sex	Grade	Washington Team	Washington Self-Contained	Longfellow Self-Contained
Girls	3	4.07	4.09	4.58
	4	5.45	4.65	5.08
	5	5.78	6.08	6.85
	6	7.36	7.83	7.62
Boys	3	3.62	3.49	4.31
	4	4.29	4.52	5.19
	5	5.83	5.43	6.32
	6	7.21	6.80	8.16

The Analysis of Variance division was 2x4 for two sexes and four grades, and separate analyses were computed on each subject. These F-ratios are given below in Table 81.

TABLE 81

F-Ratios of Iowa Test of Basic Skills Scores for Third-Sixth Grades by Subject Area and Source

Academic Year 1962-63

	Subject Area					
Source	Vocabulary	Reading	Total Language	Total Work Study Skills	Total Arithmetic	Composite
Organization	7.51***	4.29*	6.61**	4.20*	10.17***	6.78**
Grade	69.20***	51.51***	67.24***	108.02***	137.65***	92.60***
Sex	1.58	1.94	15.33***	3.67	3.81	5.18*
Org x Grade	-	-	-	-	-	-
Org x Sex	1.12	1.02	-	1.65	-	1.13
Grade x Sex	-	-	-	-	-	-
Org x Grade x Sex	1.39	1.10	-	-	-	1.03

$^{*}p < .05$
$^{**}p < .01$
$^{***}p < .001$

Organization was significant in all subjects. Table 82, below, shows the mean grade placements by organization.

TABLE 82

Iowa Test of Basic Skills: Overall Mean Grade Placement by Subject and Organization; Third-Sixth Grades

Academic Year 1962-63

Subject	Washington Team	Washington Self-Contained	Longfellow Self-Contained	Duncan Range WT-WSC	Duncan Range WT-LSC	Duncan Range WSC-LSC
Vocabulary	5.68	5.36	6.38	.48	.44	.47
Reading	5.53	5.27	6.09	.51	.46	.49
Total Language	5.46	5.50	6.35	.50	.45	.48
Total Work Study Skills	5.29	5.32	5.87	.37	.35	.34
Total Arithmetic	5.06	5.27	5.88	.33	.31	.30
Overall Composite	5.40	5.34	6.11	.40	.37	.39

In these scores there is a marked shift between the first and second grade scores: Longfellow Self-Contained is always significantly higher than both other organizations, while there are no significant differences in scores between Washington Team and Washington Self-Contained.

Grade, of course, showed significance due to the higher achievement score from one grade to the next.

Sex was significant in composite and highly significant in total language. These two scores are shown by grade and sex in Tables 83 and 84.

TABLE 83

Iowa Test of Basic Skills: Total Language Mean Grade Placement by Grade and Sex; Third-Sixth Grades

Academic Year 1962-63

	Sex	
Grade	Girls	Boys
3	4.35	3.73
4	5.40	4.82
5	6.86	6.02
6	8.04	7.39

TABLE 84

Iowa Test of Basic Skills: Composite Mean Grade Placement by Grade and Sex; Third-Sixth Grades

Academic Year 1962-63

	Sex	
Grade	Girls	Boys
3	4.29	3.83
4	5.06	4.76
5	6.29	5.95
6	7.60	7.58

VI. ACHIEVEMENT: DESCRIPTIVE ANALYSIS

The achievement post-test scores provide a basis for comparison of the three organizations in the first and second years. (In this comparison the first year first graders may be included, although they were eliminated from the covariance analysis because no measure of I.Q., one of the covariates, was available.

Table 85 below shows the basis for the comparison of the scores from the California Achievement Tests with the scores from the Iowa Test of Basic Skills, for grades 3-6.

TABLE 85

Tests Used in Descriptive Analysis of Student Achievement

Academic Years 1961-63

Descriptive Analysis Heading	First Year's Test	Second Year's Test
Reading*	CAT Reading Score	CAT Reading Score (Grades 1 and 2) ITBS: Average of Reading and Vocabulary Scores (Grades 3-6)
Arithmetic	CAT Arithmetic Score (All Grades)	CAT Arithmetic Score (Grades 1 and 2) ITBS: Total Arithmetic Score (Grades 3-6)
Language	CAT Language Score (All Grades)	CAT Language Score (Grades 1 and 2) ITBS: Total Language Score (Grades 3-6)
Overall Average	CAT Total Achievement Score (All Grades)	CAT Total Achievement Score (Grades 1 and 2) ITBS: Overall Composite Score (Grades 3-6)

*The CAT Reading Score includes a vocabulary and a comprehension score; hence, the ITBS vocabulary and reading scores were averaged to permit a score more reliable for comparison than the ITBS reading score alone would have been.

TABLE 86

Relative Differences in Mean Grade Placement Between Organizations, on Achievement Test Scores

Academic Years 1961-63

		1st Year: 1961-1962			2nd Year: 1962-1963		
Grade	Subject Area	Team	WSC	LSC	Team	WSC	LSC
1	Reading	0.70	0.00	0.64	0.63	0.00	0.53
	Arithmetic	0.45	0.00	0.30	0.48	0.00	0.09
	Language	0.69	0.00	0.53	1.06	0.00	0.95
	Total Achievement	0.61	0.00	0.49	0.72	0.00	0.52
2	Reading	0.09	0.00	0.46	0.74	0.00	0.84
	Arithmetic	0.00	0.11	0.28	0.46	0.00	0.30
	Language	0.15	0.00	0.42	0.85	0.00	0.82
	Total Achievement	0.05	0.00	0.35	0.68	0.00	0.65
3	Reading	0.00	0.27	0.77	0.00	0.06	0.66
	Arithmetic	0.00	0.22	1.26	0.02	0.00	0.70
	Language	0.04	0.00	0.82	0.20	0.00	0.86
	Total Achievement	0.00	0.15	0.94	0.03	0.00	0.67
4	Reading	0.07	0.00	0.52	0.44	0.00	0.50
	Arithmetic	0.00	0.36	0.41	0.07	0.00	0.47
	Language	0.08	0.00	0.61	0.46	0.00	0.92
	Total Achievement	0.00	0.07	0.47	0.29	0.00	0.56
5	Reading	0.00	0.43	0.29	0.37	0.00	1.16
	Arithmetic	0.00	0.62	0.22	0.00	0.33	0.66
	Language	0.00	0.78	0.12	0.00	0.31	1.08
	Total Achievement	0.00	0.58	0.18	0.02	0.00	0.78
6	Reading	0.00	0.63	1.15	0.46	0.00	0.85
	Arithmetic	0.00	2.52	2.27	0.00	0.59	0.97
	Language	0.00	0.91	0.96	0.00	0.49	0.75
	Total Achievement	0.00	1.36	1.46	0.00	0.08	0.58

Table 86, above, shows the differences in post-test mean grade placements between the lowest score among the three organizations and the other two higher scores. The organization that scored lowest on a particular grade placement mean has been assigned a zero; and the other two organizations received their scores by subtracting from their grade placement means the grade placement mean of the lowest organization.

Table 86 can be read in two ways: First, by comparing any grade for year one and year two, the reader can note whether any important changes in organizational achievement took place between the first and second years. Second, by comparing any of the first-year grades (except grade 6) with the following second-year grades, the reader can note generally how any single class did over the two year period. Obviously, these comparisons across the full two-year study must be tempered by the fact that some changes in class membership occurred between years.

When comparing a single grade for year one and year two, the following points should be noted:

1. Grade one changed extremely little from year one to year two, although a slight improvement in the position of the Team is evident in the second year. For both years, Washington Self-Contained scored lowest in all areas, with the Team scoring slightly ahead of Longfellow Self-Contained.

2. Grade two shows a definite change from year one to year two. Whereas Team and Washington Self-Contained were approximately equal during the first year, with Longfellow Self-Contained three to four months ahead in achievement, the second year found the Team achieving on a level that slightly

surpassed the Longfellow Self-Contained organization. The second year resulted in Washington Self-Contained falling behind to a greater extent in achievement.

3. Grade three, year one, shows Team somewhat behind Washington Self-Contained and also both Washington School organizations definitely eclipsed by the Longfellow Self-Contained class. Year two finds Longfellow Self-Contained still ahead, although by a lesser amount, while the Team has caught up with and slightly passed Washington Self-Contained.

4. Grade four has Longfellow Self-Contained ahead in achievement first* and second years. Team and Washington Self-Contained were about equal in achievement for year one but Team was noticeably ahead of Washington Self-Contained for the second year.

5. The first year, Washington Self-Contained achieved greatest in grade five, followed by Longfellow Self-Contained. In the second year for this grade, Longfellow Self-Contained definitely achieved the greatest amount, while Team and Washington Self-Contained scored about the same.

6. Grade six shows Team behind both years, the first year by a very large amount. Longfellow Self-Contained was only slightly ahead of Washington Self-Contained during the first year, but increased this margin during the second year.

7. Overall, comparing year two position with year one position, the Team did well in Grades 1, 2, and 4, while Longfellow Self-Contained

---

* This illustrates the difference between raw and adjusted means, in so far as Longfellow fell below both organizations in adjusted mean scores; see Figure 23.

did well in Grades 3, 5, and 6 (with the Team showing slight improvement in Grade 3.)

When following a particular class as it moves from the first year to the second year, these points should be noted:

1. Grade one, year one, to grade two, year two, shows almost no change of any kind in organizational achievement.

2. Grade two, first year, to grade three, second year, points up a gain for Longfellow Self-Contained in achievement relative to the other two organizations while Team and Washington Self-Contained changed very little.

3. Grade three, first year, to grade four, second year, shows a sizeable gain for the Team and a marked loss for Longfellow Self-Contained.

4. Grade four, year one, to grade five, year two, witnessed a gain for Longfellow Self-Contained with no other noticeable change.

5. Fifth grade, first year, to sixth grade, second year, showed a relative loss for Washington Self-Contained and a relative gain for Longfellow Self-Contained.

VII. ACHIEVEMENT AND PERSONALITY: CONCLUSIONS

It is difficult to make a direct comparison between the results of the first and second year studies, because of the use of the Analysis of Covariance in the former, and the Analysis of Variance in the latter. The main effects and interactions of the first year analyses were based on adjusted means, and primarily reflected the gain between pre- and post-test scores, * while those of the second study were based entirely on post-test scores.

Nevertheless, it is clear in general that the different classroom organizations produced much greater differences in student achievement scores than in measures of student personality. Only in personal adjustment scores of the first year was there a significant organizational effect: Washington Team was lower than the self-contained organizations in overall gain.

One might note, in reference to these scores (Table 50, page 134), that Washington Team had a lower overall mean for post-test than the self-contained organizations. Although this fact alone would not necessarily have caused a significant main effect in the covariance analysis, it can be observed that this difference did not exist in the overall mean of the second year post-tests. These were: Washington Team 33.234, Washington Self-Contained 32.500, and Longfellow Self-Contained 33.198, ** for personal adjustment on the primary level; and, respectively, 51.365, 52.883, and 52.835, on the intermediate level.*

---

* It must be remembered that I. Q. was also a covariate in the first year achievement test analyses. However, as has been pointed out I. Q. did not vary greatly.

** The second year primary level scores were not transformed to the intermediate level scale, as in the first year study.

The organization by grade interaction in the first year personality results was caused mainly by the fact that the first grade in the Team had very low gains in comparison with the first grade of Washington Self-Contained which had high gains. In fact, some of the Washington Team first grade scores dropped between pre- and post-tests, as is evident in Table 46.

In achievement, however, there were significant differences between organizations in all scores except first year arithmetic. These differences were presented in the previous descriptive section, Table 86. The first grade of the Team was higher than the first grades of the self-contained organizations in both years, and the second grade of the Team was higher in the second year. Third through fifth grades in the Team moved ahead of Washington School Self-Contained in the second year. In all grades, except the first two, Longfellow Self-Contained classes achieved higher than the other organizations in both years, except for the fifth grades of the first year, when Washington Self-Contained was highest.

These results can be seen in the graphs of Figures 25 and 26 of the total achievement and composite mean grade placements for the post-tests of both years.

The organization by grade interaction of the first year was caused mainly by the drop from very high achievement gains in the third grade of

---

* These means are based on Tables 51 and 52.

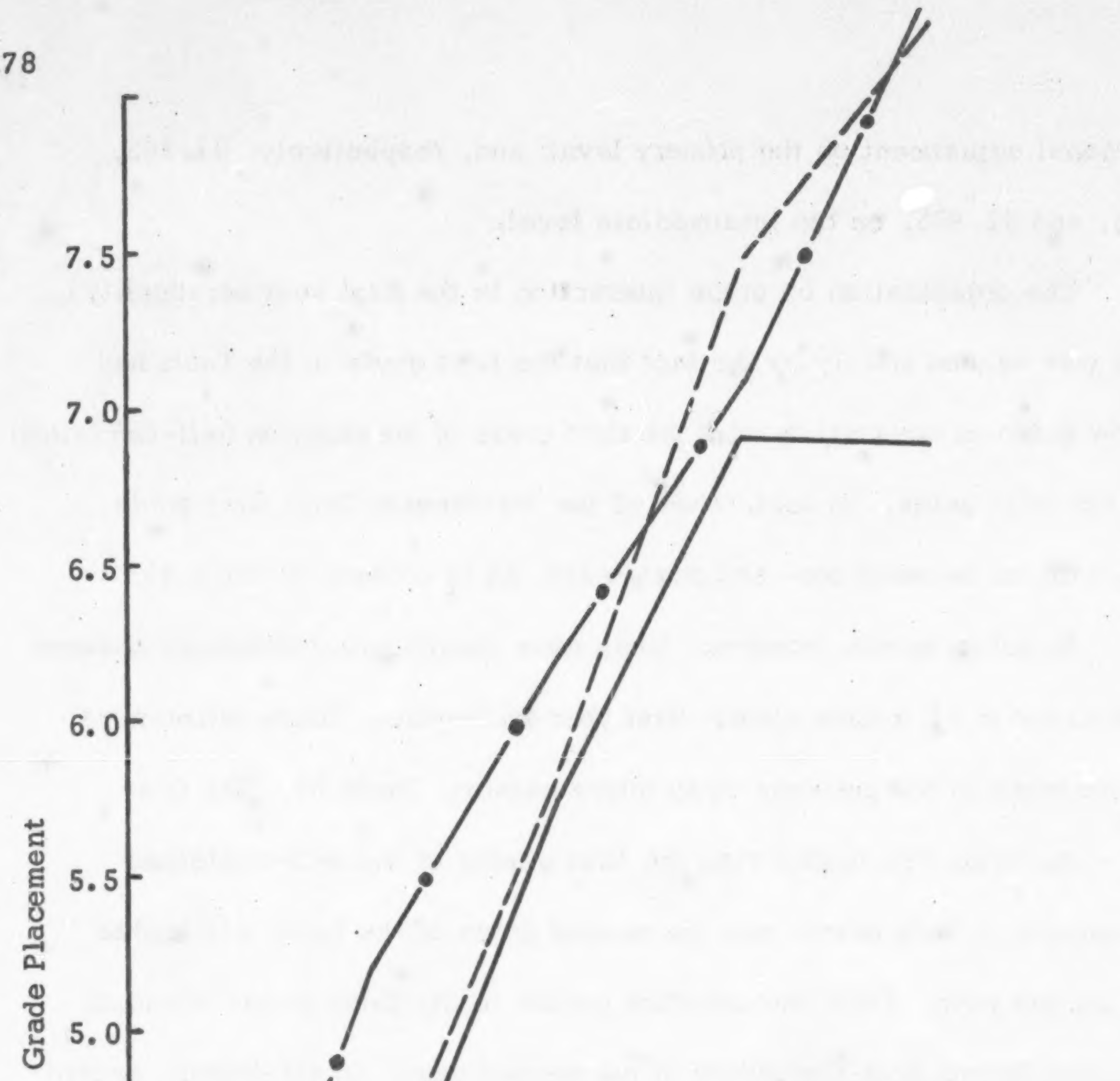

Figure 25: California Achievement Test: Mean grade placements for total achievement post-test by grade and organization; 1961-1962.

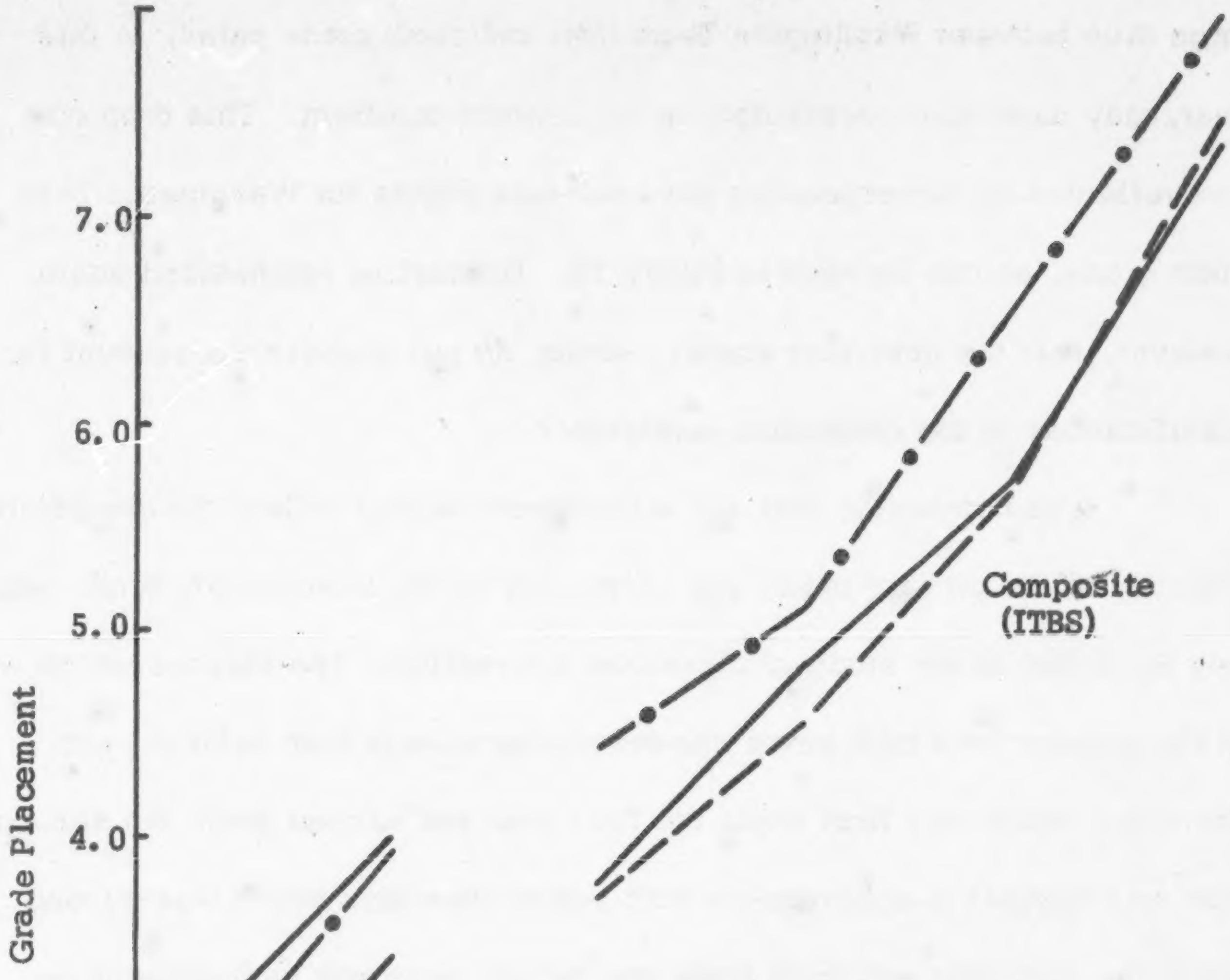

Figure 26: California Achievement Test and Iowa Test of Basic Skills: Mean grade placements for total achievement (CAT) and composite (ITBS) post-tests by grade and organization; 1962-1963.

Longfellow School to very low achievement gains in the fourth grade. A large drop between Washington Team fifth and sixth grade gains, in this year, may have also contributed to the interaction effect. This drop off was reflected by corresponding low post-test scores for Washington Team sixth grade, as can be seen in Figure 25. (It must be emphasized again, however, that the post-test scores, alone, do not necessarily account for significances in the covariance analysis.)

It is noteworthy that the achievement scores reflect the comparative stability of the primary team, and instability of the intermediate team, which was indicated in the study of classroom interaction. The classes which were in the primary team both years showed comparatively high achievement: The class which was first grade the first year and second grade the second year was highest in achievement both years; the class which was second grade the first year and third grade the second year was high in achievement, and only suffers in comparison with Longfellow Self-Contained third grade classes, which were unusually high in achievement, as has been previously observed.

However, the Intermediate Team can be seen to have improved in achievement from the first to second year, in comparison to Washington School Self-Contained classes, although it remained consistently low in comparison to Longfellow Self-Contained Classes.

The class which was Primary Team third grade in the first year and Intermediate Team fourth grade in the second year, was highest in comparative achievement of the three Intermediate Team grades of the second year, as can be seen in Table 86 of the Descriptive Analysis.

PART FOUR

SIDE STUDIES

PRECEDING PAGE BLANK-NOT FILMED

I. TEACHER AWARENESS OF STUDENT CHARACTERISTICS

A. Purpose

The purpose of this study was to determine if any difference exists between team and self-contained teachers relative to their awareness of factual information about their students. Direct comparisons could be made between the teacher's estimate and the actual information as found in the students' cumulative records.

B. Procedure and Administration

This study used a nine-item questionnaire* that served as a test of the teacher's knowledge about her students. Teachers filled out a separate questionnaire for each of a selected number of their pupils. The responses of a teacher for a particular student were compared to the cumulative record information for that student. The first eight questions were analyzed; the ninth was added to avoid "penalizing" a teacher because her awareness was more current than the cumulative record.

The students included in the study were randomly chosen from the classes of the 20 participating teachers for the first year of the Project. Fifteen students were chosen for each teacher. The questionnaire was administered twice. It was realized that some teachers might be prompted by the first administration to become better informed on their pupils, but it was assumed that this was as likely to occur in each of the organizations.

---

* The questionnaire and the instructions that accompanied it may be found in Appendix C.

1) First Administration: April 2, 1962

Three organizations (Washington Team, and Washington and Longfellow Self-Contained) were included in the first administration. Each teacher responded individually on 15 pupils. With eight team teachers and six teachers from each of the two self-contained organizations, a total of 300 students were involved.

2) Second Administration: April 16, 1962

The second administration included only the 14 teachers, Team and Self-Contained, at Washington School; the same pupils were used as in the first administration. The four Primary Team teachers were allowed to respond as a group, pooling their information on their students; the same provision was made for the Intermediate Team. Longfellow School was not included since there was no significant difference between the two self-contained organizations in the first administration. Furthermore, 15 students from each of the two teams were randomly omitted from the analysis; this reduced the Team students from 120 to 90, the same number as available for Washington Self-Contained.

C. Statistical Analysis

1) Questions 1, 2, and 4-7

For each response on the first seven questions (except question 3), the teacher was given an absolute error score (the difference, regardless of sign, between teacher response and actual information as recorded in the student's cumulative record). Then all the absolute error scores (300 on the first administration and 180 on the second) for a particular question, were pooled and converted to T-scores (mean 50 and standard deviation 10). Therefore, the lower the resulting T-score for error, the more aware the teacher was of her student's actual characteristic on that item (question).

On the first administration, the 90 T-scores for error available on each teacher (15 pupils times 6 questions) were averaged to yield a single composite-error score for each teacher. The 20 composite-error scores resulting were subjected to a 2 x 3 ANOVA (level by organization).

On the second administration, a different procedure was followed. In this case, the 6 T-scores for error available on each pupil were averaged to yield a single composite-error score on each pupil. The 180 composite-error scores that resulted were analyzed in a 2 x 2 ANOVA (level by organization) with 45 observations per cell.

2) Question 3

Question 3 elicited a "yes" or "no" response. With no error scale, the responses to it were tabulated as either right or wrong. No statistical analysis of the results was undertaken, but a descriptive analysis was made.

3) Question 8

Question 8 required a teacher to estimate how her students had been rated by peers on a scoiometric rating scale. This question applied only to the fourth, fifth, and sixth grades, i.e., the intermediate level, and was analyzed only for the first administration. An absolute error score was obtained by comparing the teacher's judgment for each student to a composite rating made previously by the student's classmates; these absolute error scores were summed for each teacher. This resulted in ten composite-error scores (four Intermediate Team teachers and six intermediate level teachers in the two self-contained organizations) which were subjected to an ANOVA, one-way classification.

D. Results

1) Questions 1, 2, and 4-7

The self-contained teachers were somewhat more aware of their students' characteristics than the Team teachers (especially the interns) on the first administration; however, on the second administration when the Team teachers pooled their opinions, this situation was reversed. In neither case was the difference significant. Table 87 gives the results of the two ANOVA's on the six questions analyzed as a group.

TABLE 87

F-Ratios of Teacher Awareness Responses (Questions 1, 2, and 4-7) for First and Second Administrations by Source

Academic Year 1961-62

	First Administration		Second Administration	
Source	df	F-Ratio	df	F-Ratio
Organization (0)	2	1.84	1	-
Instructional Levels (L)	1	-	1	1.42
O x L	2	-	1	3.22
Error	14		176	

Note: The error term for the first administration is teachers within group; for the second administration, the error term is the within class variation of the teachers.

2) Question 3

The percents of correct to total responses on Question 3 are given in Table 88. In Table 89, the Team master-intern breakdown for the First Administration is given.

TABLE 88

Percents of Correct to Total Responses on Question 3 of the Teacher Awareness Questionnaire for First and Second Administrations by Organization and Level*

Academic Year 1961-62

	First Administration			Second Administration	
Level	WT	WSC	LSC	Team	WSC
Primary	88.3%	88.9%	95.6%	80.0%	93.3%
Intermediate	86.7%	91.1%	91.1%	91.1%	88.9%

* All percents are based on 45 observation except those for the Primary and Intermediate Teams, First Administration, which are based on 60.

TABLE 89

Percents of Correct to Total Responses on Question 3 of the Teacher Awareness Questionnaire for First Administration by Type of Team Teacher and Level*

Academic Year 1961-62

Level	Interns	Masters
Primary	80.0%	96.7%
Intermediate	76.7%	96.7%

* All percents are based on 30 observations.

3) Question 8

The results of the ANOVA on question 8 (first administration) are given below in Table 90; organization was not significant.

TABLE 90

F-Ratios of Teacher Awareness Responses (Question 8)
for First Administration by Source

Academic Year 1961-62

Source	df	F-Ratio
Organization	2	2.12
Error	7	

E. Discussion and Conclusions

1) Questions 1, 2, 4-7

The results of the analyses of variance on the 6 questions give no basis for rejecting the null hypothesis, that is, there is no significant difference in teacher awareness of student characteristics between the teachers of the team and the self-contained organizations. This was true whether team teachers responded individually (first administration) or as a group (second administration).

2) Question 3

The percents of correct to total responses for question 3 do not indicate any differences between the self-contained classroom teachers and the master teachers of the team. The difference between master teachers and interns, however, is quite large. This is not unexpected as the interns had only been on the teams slightly over two months when the questionnaires were given. It is interesting to note that the Primary Team teachers did slightly poorer on the second administration than the first, even though they were allowed to pool their opinions on the second administration.

3) Question 8

The data obtained on this question indicate that there is no significant difference in teacher awareness of peer sociometric ratings between teachers of the team and self-contained teachers.

II. DISTRIBUTION OF ABSENCE

The absenteeism of 365 students involved in the first year of the Project was analyzed to see if there were any significant differences between the organizations. There were 109 students in the Team, and 256 in self-contained classrooms: (110 at Washington and 146 at Longfellow.)

The mean number of half days absent by organization and instructional level appear in Table 91.

TABLE 91

Mean Number of Half-Days Absent by Organization and Instructional Level

Academic Year 1961-62

Level	WT	WSC	LSC
Primary	19.73	25.29	20.72
Intermediate	18.13	20.69	16.81

Assuming that the overall average of half days absent across organizations would be the expected frequency for each cell, a chi-square test was computed on observed frequencies. This test turned out to be non-significant. Thus the type of organization does not affect the amount of absenteeism any more than would be expected by chance. *

---

* Noall and Jensen (118) report better attendance for an eighth grade team class than for seventh and ninth grade control classes, but the small difference between the attendance figures between the eighth and ninth grade classes implies that the only statistically significant difference was between the seventh and eighth grades.

III. STUDENT DISCIPLINE

A. Purpose of the Study

The purpose of this study was to investigate the differences in frequency and seriousness of discipline problems between the self-contained classroom and the team teaching organizations.

B. Model Used for Investigation

Two factors concerning discipline were investigated: (1) the frequency of undesirable incidents; and (2) the seriousness of these incidents. The observation model used consisted of five categories, as follows:

Category 1

1. Speaking without permission but doing so relative to the particular lesson under study.
2. Failing to get down to business promptly when the teacher or the class is waiting.
3. Whispering without permission in the non-participating group.

Category 2

1. Attracting attention in an undesirable manner, such as continually dropping a pencil.
2. Playing with a pencil or other object.
3. Leaving seat without permission.
4. Showing lack of attention corrected by the teacher.
5. Talking back to the teacher in a defensive way.

Category 3

1. Making a wise remark.
2. Tipping a chair or disrupting furniture for the purpose of causing a disturbance.
3. Period of general confusion and lack of attention in the classroom.
4. Minor horseplay between students.

Category 4

1. Passive response to teacher's direction.
2. Persistence in lack of attention.
3. Extended horseplay with a reluctance to stop.

Category 5

1. Talking back to the teacher in a sarcastic way.
2. Refusing to obey the teacher when given a direct command.
3. Fighting, or conducting in a hostile verbal exchange with the teacher or other students.
4. Malicious physical action directed at the teacher.

Directions for Use of the Discipline Model

The observation should cover a period of 30 minutes. Begin observing when the class has settled down and the teacher has control.

Each incident should be tallied once, that is, in only one category.

Judgment is required as to the seriousness of the incident. However, each incident should be tallied into the category it fits or most nearly fits.

During each week of observation (following the training period) observe together at least one class in order to get reliability checks during the actual observation.

The categories for the model are mutually exclusive and exhaustive. The items in each category were judged to be of essentially equal seriousness by the investigators. The order of seriousness of categories ranged from the least serious, Category 1 to the most serious, Category 5. The category scale was considered an ordinal but not an interval scale. That is, Category 3 was considered more serious than Category 2 but the difference between seriousness of Categories 1 and 2 might or might not have been the same as the difference between Categories 2 and 3.

C. Observers and Data Collection

Two former public school teachers (women) were trained as observers on the discipline model. The training period took place during the week of November 26, 1962. The data collection period consisted of four weeks: December 3 through December 14, 1962, and January 7 through January 18, 1963.

The duration of each observation period was thirty minutes beginning at the onset of class. The data gathered consisted of the frequencies tabulated in each category of the model. Observers were situated near the rear of the room.

The observers jointly observed ten periods during the training period. They observed the same period, making independent ratings. From the joint observations a measure of observer agreement was obtained.

The measure of agreement was a percentage score computed by the formula:

$$\%\text{Agreement} = \frac{\text{No. of tallies in agreement}}{\text{Avg. number of tallies}} \times 100$$

It should be noted that the denominator consists of an average rather than a total number. Unlike models such as the interaction models, the total number of tallies for a thirty minutes period may vary between the observers, hence an average is used. The results of the ten training periods showed an average observer agreement of 74% with a range of 50% to 100%. During the training period the observers discussed each period following the observation in an effort to eliminate future differences in definition and interpretation. Although all attempts were made to keep the use of the model as objective as possible, some observer judgment was involved.

In addition to the training period, observer agreement checks were made during the data gathering period. The percentage of agreement during the data gathering period maintained reliability.

The observations were made on all three organizations within the study. Self-contained teachers were observed four times and each team teacher three times. This was done in order to balance out the number of observations for each group and level. The observations took place during reading, arithmetic, and social studies classes with each teacher observed at least once in each subject area. However, there were a few exceptions to this due to the team structure. Team teachers did not all teach all subjects, and in one case literature was included as a subject to be observed.

D. Statistical Analysis

The frequencies tabulated in each category, when broken down by groups for comparisons, did not follow a normal curve. There is no reason to believe that they should. The frequencies for each category are concentrated near the low end of the natural number system, with a few extreme scores going the other way. This concentration with the extreme scores make the distributions skewed to the left, with a cutoff at zero, and extended to the right.

In order to test statistically for differences between the organizations under investigation, non-parametric methods were used. The frequencies of the categories of the model were different. The number of frequencies decreased markedly as the seriousness of the incidents which made up the category increased. This was the expected and actual situation. Each category was analyzed separately.

Categories 1 and 2 were analyzed using the Kruskal-Wallis One-Way Analysis of Variance by Ranks. * This is a technique which ranks all of the observations from 1 to N in order of smallest to largest. The assumption is that the variable under investigation has an underlying continuous

---

* Siegel, Sidney, Nonparametric Statistics, pp. 184-193.

$$H = \frac{12}{N\,(N+1)} \sum_{j=1}^{k} \frac{Ri^{2}}{Nj} - 3\,(N+1)$$

k = number of samples
Nj = number of cases in jth sample
N = total frequency
Rj = sum of ranks in jth sample

distribution and that at least ordinal measurement of that variable is possible.

The computation of the statistic "H" (the statistic used in the Kruskal-Wallis Test) determines whether the sums of the ranks of the various samples are so disparate that the samples are unlikely to have been drawn from the same population. H is distributed as chi-square with df = k-1.

The three major organizations were compared for each of the two levels. The master teachers and interns were compared including both Teams. The self-contained and team organizations of Washington School were compared for each of the two levels. Finally the teams were compared. Each of these comparisons was done separately for Categories 1 and 2.

The frequencies of Categories 3, 4, and 5 dropped markedly from those of Categories 1 and 2. Chi-square tests on total frequency were used between the three major groups for each of the two levels on Categories 3 and 4. The remainder of the comparisons and the frequencies of Category 5 are descriptively reported.

E. Results

The H values of the Kruskal-Wallis test are given in Table 92 by the various comparison groups.

TABLE 92

H-Values of the Kruskal-Wallis Test by Category and Comparison Group

	Category	
Comparison Group	1	2
LSC-WSC-WT Primary	5.03	1.95
LSC-WSC-WT Intermediate	2.64	2.70
Master Teachers-Interns	2.90	4.08*
WSC-WT Primary	1.61	1.84
WSC-WT Intermediate	-	1.84
PT-IT	-	-

*$p < .05$

Category 1 approached significance for the comparison of the three organizations on the primary level. Washington Self-Contained ranked high and Longfellow Self-Contained low with the Washington Team falling almost right in the middle on the total sum of ranks. On the significant value in Category 2, master teachers-interns, the interns were high.

Although only one comparison was significant, the relative position of the sums of ranks for the three organizations are worth noting. They are given in the following diagram:

		WT	WSC	LSC
Primary	Category 1	medium	high	low
	Category 2	low	high	medium
Intermediate	Category 1	high	medium	low
	Category 2	medium	high	low

The pattern is definitely one of the extreme difference being between the two self-contained control groups. This indicates that the particular school, with its individual student body and general discipline attitude, is a more relevant factor than the class structure.

The master teachers were lower than the interns for both categories and in the one case significantly so.

The total frequencies and the resulting Chi-square values for Categories 3 and 4 are given in Table 93.

TABLE 93

Frequencies and Chi-Square Values
by Level, Group, and Category

Category and Level	WT	WSC	LSC	X^2
3 - Primary	11	22	8	7.93*
3 - Intermediate	34	30	10	13.39**
4 - Primary	1	3	3	1.16
4 - Intermediate	15	15	2	10.53**

*$p < .05$
**$p < .01$

The significance of the primary level, Category 3, shows Washington Self-Contained to be high with Longfellow Self-Contained and Washington Team very close in total frequency. However, on the intermediate level for both categories the two Washington organizations were very close with the Longfellow Self-Contained low. The frequencies of master teachers for the two categories were 10 and 5 and those for interns 35 and 11, respectively.

The frequencies for Category 5 showed none at Longfellow for either level, 3 at the Washington Self-Contained and 5 at the Team all at the intermediate level. There were none at Washington for the primary level. Essentially no generalizations can be drawn from this since the few incidences may have been an interaction with a particular unruly student.

IV. SOCIAL STRUCTURE

A. Introduction

The social structure of a group is defined as the aggregate of all interpersonal relations between pairs of individuals in the group (Katz and Proctor, 88). It was assumed for purposes of this study, that change in the social structure is a time-dependent, stochastic process. This essentially means that the social structure state at any one time depends on some chance element but that in a sequence of states, the state at any one time is dependent upon the state(s) at a previous time or times.

B. Data Getherine Instrument and Schedule

The data gathering instrument used in this study consisted of a three-point "like to play with" sociometric rating scale. The three points of the scale were: 1) especially like to play with (E), 2) sometimes like to play with (S), and 3) least like to play with (L).

Each student rated every other student on the three point scale. Thus, the social relationship existing between all pairs of students within a group comprised the social structure of the group.

The fourth, fifth, and sixth grade students of the 1961-62 year were included in this study. A total of 185 students across the three organizations participated. Thus, the study included only the intermediate level. The distribution of students by organization and grade is given in Table 94.

There were three data gathering sessions. These occurred at two month intervals during February, April, and June of 1961. The rating scale was administered to the entire group at the same time. A short pilot study was previously conducted to check for an effect of method of administration i.e., personal interview or group administration. An ANOVA showed no significant method effect and hence the group method was assumed not to have biased the results.

TABLE 94

Distribution of Students for Social Structure Study by
Organization and Grade

Second Semester 1961-62

Grade	Washington Team	Washington Self-Contained	Longfellow Self-Contained
4		17	23
5	52	19	26
6		16	32

The tabulation of the data consisted of placing the student's response ratings into an n x n array for each group where n equalled the number of students in a group. These ratings were then grouped into a dyad for each pair. There were nine possible dyads, namely, EE, ES, EL, SE, SS, SL, LE, LS, and LL. The frequency of each dyad occurring was then tabulated. The comparison of the social state of a particular group at two times, say T_1 and T_2, was then represented by a 9 x 9 array of the frequencies of the changes (or stability on the diagonal) in the dyads. One of the times was represented horizontally and the other vertically.

C. The Markov Chain

The purpose of this tabulation was to apply the theory of Markov Chains to the analysis. A Markov Chain may be considered as a chain of matrices in which any particular matrix or state is dependent upon one or more of the previous states. The assumptions for a Markov Chain may be summarized as follows: The process begins in some particular state and each state is one of a finite number of possible states. There exist numbers p_{ij} called transition probabilities which represent the probability of the outcome of a specific state, say a_i, given that state a_j occurred on the preceding state. The probability of a given state existing at any given time is not independent of previous states.

A Markov Chain is defined to be of order one if the state at one time is dependent only on the previous time. This generalizes to order "k" if the state at any one time depends on the "k" previous states.

Katz and Proctor (88) demonstrated that change in this process based on four-month intervals conforms to a Markov Chain of order one. They found that the transition probabilities were constant from one pair of time points to the next (based on four- and two-month intervals). Having determined these characteristics, they were able to estimate the transition probabilities.

This study is primarily concerned with the transition probabilities, and specifically, with whether the type of classroom organization has a significant effect on the transition probabilities.

D. Statistical Analysis

The analysis was divided into two parts. The first part includes the steps involved in determining whether the process conformed to a Markov Chain of order one. The procedure is given in Table 95. Chi-square tests were used in all cases. The sequence of steps is important since previous hypotheses must hold.

The second part of the analysis dealt with the direction and amount of change in the relationships, regardless of the original type of relationship. There are 15 types of transition as shown in Table 96.

Chi-square tests were again used in Part II of the analysis. A summary of the procedure of Part II is given in Table 97.

TABLE 95

Summary of the Procedure of Part I of the Analysis Based on Chi-Square Tests

What Tested	Theoretical Frequency	d.f.	Interpretation of X^2 Test
Dependence of States from T_j to T_{j+1}	$f_{ij} = \frac{n_{i.}n_{.j}}{n_{..}}$	$(k - 1)^2$	X^2 significant indicates dependence
Conformity to a Markov Chain	$f_{ijt} = \frac{n_{ij}}{2}$	$(k - 1)(T - 1)$	X^2 not significant indicates Markovity
Order of the Markov Chain	$f_{ij} = \frac{n_{i.}n_{.y}}{n_{..}}$	$(I - 1)(J - 1)$	X^2 not significant indicates order one
Similarity of Transitions between Classrooms	marginal totals applied to $T_j - T_{j+1}$	$(R - 1)(J - 1)$	X^2 not significant indicates similarity

Where:

n_{ij} is observed frequency in each cell, dot indicates summation over the missing subscript;

k is the number of rows or columns of the transition matrix;

T is the number of transitions;

I is the number of rows; J is number of columns; and

R is the number of classrooms

TABLE 96

Fifteen Transition Categories Indicating Direction and Amount of Change in Relationships between Two Points in Time

Category	Description	Transitions From ______ to ______
1	one partner goes up 1	ES-EE, EL-ES, SE-EE, SS-ES, SS-SE, SL-EL, SL-SS, LE-SE, SL-SS, LS-LE, LL-SL, LL-LS
2	one partner goes up 2	EL-EE, SL-SE, LE-EE, LS-ES, LL-EL, LL-LE
3	both partners go up 1	SS-EE, SL-ES, LS-SE, LL-SS
4	both partners go up 2	LL-EE
5	one partner goes up 1, one partner goes up 2	LL-ES, LL-SE, LS-EE, SL-EE
6	one partner goes down 1	EE-ES, EE-SE, ES-EL, ES-SS, EL-SL, SE-SS, SE-LE, SS-SL, SS-LS, SL-LL, LE-LS, LS-LL
7	one partner goes down 2	EE-EL, EE-LE, ES-LS, SS-LL, SE-SL, LS-LL
8	both partners go down 1	EE-SS, ES-SL, SE-LS, SS-LL
9	both partners go down 2	EE-LL
10	one partner goes down 1, one partner goes down 2	EE-SL, EE-LS, ES-LL, SE-LL
11	one partner goes up 1, one partner goes down 1	EE-SE, EL-SS, SE-ES, SS-EL, SS-LE, SL-LS, LE-SS, LS-SL
12	one partner goes up 2, one partner goes down 1	EL-SE, SL-LE, LE-ES, LS-EL
13	one partner goes up 1, one partner goes down 2	ES-LE, EL-LS, SE-EL, LE-SL
14	one partner goes down 2	EL-LE, LE-EL
15	no change	EE-EE, ES-ES, EL-EL, SE-SE, SS-SS, SL-SL, LE-LE, LS-LS, LL-LL

TABLE 97

Summary of the Procedure of Part II
of the Analysis Based on Chi-Square Tests

What Tested	Theoretical Frequency	df	Interpretation
Similarity of Transitions between Times	$f_{ct} = \frac{NC.N._{t}}{N..}$	$(C-1)(T-1)$	X^2 not significant indicates similarity
Similarity of Transition between Classrooms; Organizations	$f_{ctr} = \frac{N_{ct.}N_{.tr}}{N.t.}$	$(C-1)(T-1)$ $(R-1)$	X^2 not significant indicates similarity

(C is number of transition categories)

E. Results

The reader is reminded that six self-contained classrooms and one team organization were observed three times each. This accounts for the seemingly large number of comparisons discussed in this section. The hypotheses are discussed in the order of appearance in Tables 95 and 97. Unless otherwise stated, significance means the .05 level.

The dependence of the states from $T_1 - T_2$ and $T_2 - T_3$ was tested for all seven groups. The 14 chi-square values were significant beyond the .0005 level. Thus the state at any one time is considered dependent upon the state at the previous time.

The comparison of the $T_1 - T_2$, $T_2 - T_3$ transition matrices supported the hypothesis of Markovity in five of the seven groups. The chi-square values were significant for the sixth grade Washington Self-Contained and the Team classrooms. Further analysis of the Team showed that no row was significant.

The sixth grade, Longfellow Self-Contained was used to determine the order of the Markov Chain. The type of relationship at T_3 was tested for independence from T_1. The results of the test are given in Table 98.

TABLE 98

Chi-Square Values for Order of Dependence of Markov Chain in Longfellow School, 6th Grade

Type of Relationship at T_2	df	Chi-Square
EE	6	4.39
ES (or SE)	20	19.09
EL (or LE)	20	36.95*
SS	25	27.70
SL (or LS)	20	34.71*
LL	15	15.71

*$p < .05$

Four of the relationships (those not significant) indicated a Markov Chain of order one, while two indicated a higher order.

The results of tests of similarity of transition probabilities between each self-contained classroom and the team are shown in Table 99.

TABLE 99

Chi-Square Values Obtained from Comparison of Transition Probabilities of Self-Contained Classrooms with Team-Teaching Organization

Initial Type of Relationship	WSC_4	WSC_5	WSC_6	LSC_4	LSC_5	LSC_6
EE	5.85	2.08	10.19	2.72	1.99	6.65
ES (or SE)	9.71	15.38**	19.98**	9.88	6.57	8.42
EL (or LE)	13.24*	4.50	13.38*	7.82	14.36*	9.40
SS	15.54**	27.86***	17.23**	10.78	15.48**	18.63**
SL (or LS)	12.54*	13.78*	18.42**	18.28**	32.79***	27.58***
LL	23.57***	2.56	20.72***	41.91***	2.76	4.00

*$p < .05$
**$p < .01$
***$p < .001$

NOTE: Subscripts in the abbreviations for oeganization indicate grade.

As indicated in the table, the hypothesis of similarity of transitions did not hold for many relationships. However, when the self-contained classrooms within schools were compared, the same situation existed. These results are given in Table 100.

The results of the test of similarity of transitions between times were not entirely consistent. The distributions of types of transitions were found significantly different in three self-contained classrooms and the Team.

The similarity of transitions between self-contained classrooms were compared for all combinations within a school. The results of the chi-square tests are reported in Table 101.

Six of the 16 tests indicated similarity of transition.

TABLE 100

Chi-Square Values Obtained From Comparison of Transition Probabilities of Self-Contained Classrooms Within Schools

Initial Type of Relationship	$WSC_{4,5,6}$	$LSC_{4,5,6}$
EE	18.52*	8.68
ES (or SE)	35.20***	13.79
EL (or LE)	9.75	55.72***
SS	26.54**	13.99
SL (or LS)	6.56	22.26*
LL	21.71**	40.84***

*$p < .05$
**$p < .01$
***$p < .001$

TABLE 101

Chi-Square Values Obtained from Comparison of Distributions of Types of Transitions in Self-Contained Classrooms

Comparison Between	$T_1 - T_2$	$T_2 - T_3$
$SCW_{4,5,6}$	41.00**	59.13***
$SCL_{4,5,6}$	51.39***	29.08
$SCW_{4,5}$	6.97	24.75***
$SCW_{5,6}$	21.79**	12.99*
$SCW_{4,6}$	12.65*	36.43***
$SCL_{4,5}$	23.19**	9.71
$SCL_{5,6}$	11.20	15.23
$SCL_{4,6}$	43.30***	12.21

*$p < .05$
**$p < .01$
***$p < .001$

The results of the same type of comparison between the Team and the six self-contained classrooms are given in Table 102.

TABLE 102

Chi-Square Values Obtained From Comparison of Distributions of Types of Transitions of Team-Teaching Organization with Each Self-Contained Classroom

Classroom	$T_1 - T_2$	$T_2 - T_3$
SCW_4	20.65	32.62**
SCW_5	12.56	17.06
SCW_6	62.91***	26.69*
SCL_4	32.74**	18.22
SCL_5	20.07	12.69
SCL_6	35.38***	20.69

*$p < .05$
**$p < .01$
***$p < .001$

Seven of the twelve tests indicated similarity of transition. Thus, this proportion is greater than the proportion of the self-contained comparisons.

F. Discussion and Conclusions

The results of this study are not entirely consistent and hence any conclusions are tentative. Generally, the data indicate that the process of change in the social structure does conform to a Markov Chain and the transition probabilities are constant through time.

The two month gap may be too short to ensure first order dependence.

The transition probabilities appear to be similar between classrooms at Longfellow School and very different between classrooms at Washington School. The transition probabilities at Longfellow School seem to be similar to those in the Team. A possible factor might be class size, since the self-contained classes at Longfellow School were considerably larger than those at Washington School.

The transition probabilities for other initial relationships indicate no pattern; there was considerable variation between self-contained classrooms and the Team.

In the analysis concerned with direction and amount of change in each transition period, it was found that some classrooms show similar distributions and others do not. In comparisons between classrooms there was considerable variability.

This study indicates that there are differences between classrooms in social structure change but that the type of classroom organization is not a relevant factor.

PART FIVE

DISCUSSION AND CONCLUSIONS

PRECEDING PAGE BLANK-NOT FILMED

I. DISCUSSION

In the introductory section of this report, two major aims of the study were described: Comparison of the team-teaching and self-contained organizations; and sophistication of the instruments of educational research. Many significant results in both areas have been described in the discussions ending Parts Two, Three, and the separate studies in Part Four, of this report. The purpose of the present section is to bring together salient features of the project which are interrelated.

1. The personality scales of personal and social adjustment were not found to be sensitive to differences between the three organizations under examination. It is possible that there were no significant organizational effects on student personality. However, it should be noted that great variation within organization occurred in the personal and social adjustment scores. This variation, itself, may have been so great as to disguise between-organization effects.

2. Achievement measures, on the other hand, proved to be highly sensitive to differences between the organizations. Significant effects appeared in both years even though the specific achievement tests were changed.

3. The interaction models were sensitive both to differences between organization and grade level differences within each organization.

a. The use of tapes, however, did not appear to be as satisfactory as within-class observations. The loss of actually seeing the class sessions is undesirable and cannot be compensated for using the tapes.

b. Between the Withall and Flanders Interaction Models, the Flanders Model was felt to be more appropriate for studying classroom interaction. Frequent observations yield much data in a relatively short period of time while the sequential order of observation permits the study of the ongoing classroom situation as it takes place.

c. The descriptive analysis suggested by Flanders makes use of the sequential arrangement of the observations and teases much valuable information out of the data.

d. Certain modifications of the Flanders Model seem appropriate. Category 10, (Silence or Confusion) could be broken into its two parts; and Category 7 (Criticizes or Justifies Authority) might be separated into strong and mild criticism, as in the Withall Model. (Modifications of the cell-groupings for the descriptive analysis, made after the data is acquired, are suggested on page 85.)

4. The two sections within Washington School, the experimental team and the control self-contained organization, developed important differences in the two year period.

a. The Team began to move ahead of the self-contained classes in achievement the first year, and over-achieved the self-contained classes in almost all scores, the second year.

b. However achievement scores of the third and sixth grades within the Team levelled off in comparison to the other grades and the control group. Thus the upper grades of the teams are somewhat hindered in achievement relative to other grades in the same team. It may be reasonable to suppose that these differences would be erased by extending the teams to cover additional grades -- possibly even an entire ungraded 1-8 elementary school.

5. In general the Team did not achieve as well as the second control group, Longfellow Self Contained. However, the Primary Team first grade in the first year, and first and second grades in the second year over achieved the corresponding grades in Longfellow School. There is reason to believe the Primary Team was emerging as a strong high-achieving organization.

6. Comparison of the two self-contained control organizations shows that great variation occurred in both interaction and achievement patterns between two organizations of the same teaching structure. This variation points to the possibility of important effects of an uncontrolled teacher variable. Particularly evident are the differences between the third grades of the two organizations in both years of the study. However, other unidentified factors in addition to the teacher variable, may have contributed to these differences.

7. Both interaction and achievement results pointed to instability within the Intermediate Team. While the Primary Team had continuing experienced teachers, both the team leader and master teacher of the Intermediate Team were replaced in the course of the Project.

8. Differences between the inexperienced interns and experienced certified teachers suggested that a team made up entirely of experienced teachers might be more desirable. There were marked differences between

master and intern teachers in awareness of the students' cumulative records (see pp. 189-190), and in their disciplinary problems with students, (see figure 12) and the side study on student discipline.

II. CONCLUSIONS

The conclusions stated here are statements relative to the rejection or failure to reject the hypotheses stated in the introductory chapter. The hypotheses at that point were stated as null hypotheses. The conclusions here are stated simply in a general way. For specific results of several tests related to a particular hypothesis the reader is referred back to the results sections of the respective chapters.

The conclusions follow the order of the hypotheses.

I. General conclusions.

1. There are significant differences in classroom interaction between the team and self-contained organizations.

2. There are no significant differences in personal and social adjustment of students between the team and self-contained organizations.

3. There are significant differences in student achievement between the team and self-contained organizations.

II. Related conclusions.

1. There are no significant differences in teacher awareness of student characteristics between the teachers of the team and self-contained organizations.

2. There is no significant difference in absenteeism between the students of the team and those of the self-contained organizations.

3. There is no significant difference in the frequency of discipline infractions between the classes of the team and self-contained organizations.

4. There are no significant differences in changes in the social structure within the classes between those of the team and self-contained organizations.

APPENDIX A

INTERACTION MODELS

PRECEDING PAGE BLANK-NOT FILMED

I. WITHALL MODEL

1. Asks for Information

An act which has as its major intent the eliciting of a response which presumably may be evaluated for its accuracy, either by objective operation, general acceptance, or reference to an authority (such as the teacher or a text book).

Examples: Asks question about content of lesson, asks for report; asks for confirmation of response previously given; asks for repetition of what has been said; offers incomplete statement with the expectation that other will finish it; asks any question in such a way as to imply that there is a "right" answer; asks name of an object; asks for definition; asks for enumeration.

2. Seeks or Accepts Direction

An act implying willingness to consider suggestion or direction from other, or if suggestion of direction has already been offered, an act or statement indicating compliance.

Examples: Asks how to begin on an assigned task; asks what to do next; asks which procedure to follow; asks for volunteers; follows directions of other; agrees with suggestion or direction; indicates that direction will be followed at some future time; asks for permission for a specific act.

3. Asks for Opinion or Analysis

An act intended to elicit problem-structuring statements from other, either affective-evaluative or cognitive-interpretive.

Examples: Asks for opinion, wish, feeling, belief, or preference; asks for evaluation of behavior; requests for interpretation or explanation of some phenomena without implying that there is one "correct" answer; requests for elaboration or examples of a concept; requests for statement of relationships between concepts; non-directive leads or questions to facilitate self-exploration of other; reflection of feeling or alternate meaning of what other has said for purposes of clarifying meaning; asks for interpretation of other's personal experience (as distinguished from asking for a report of experience).

4. Listens

Five consecutive seconds or more out of any five second interval where an individual is listening, or attending to another individual is given a listening score (less than five seconds is not scored).

5. Gives Information

An act intended to convey, confirm, or inform "facts" which may be evaluated by objective operation, general acceptance, or reference to an authority.

Examples: Giving data such as names, dates, speed, capacity, etc., relevant to a topic under discussion; attempting to provide information requested by other; confirming the accuracy of other's response; denying the accuracy of other's response; gives report on what one has seen, heard, read, etc; gives repetition of what has been said; names object; gives definition; gives enumeration.

6. Gives Suggestion

An act intended to structure action or indicate alternatives for other which, at the same time, implies autonomy for other by providing more than one alternative or allowing for refusal.

Examples: Offering a procedure in a tentative way; offering two or more procedures, leaving choice to other; stating a preferred behavior without indicating that the preference holds for other; volunteers own services.

7. Gives Directions

An act intended to structure some action of other in which compliance seems to be taken for granted, or in which non-compliance would probably elicit some form of disapproval.

Examples: Calling class to attention; calling attention to some detail; getting attention of other by calling his name; routine administrative directions or orders; stating expectation of behavior to be followed; setting limits on behavior; stating consequences of behavior; granting a request; denying a request.

8. Gives Opinion

An act intended to structure or give direction to a topic under discussion by use of speaker's internal, private, or unstated criteria.

Examples: States opinion, wish, feeling, belief, or preference; makes a statement or asks a question reflecting a personal point of view; verbalizes introspective processes; gives criticism or evaluation of a behavior or concept; agrees or disagrees with opinion voiced by other.

9. Gives Analysis

An act intended to structure or give direction to a topic under discussion by reference to a frame of reference or a criterion that is explicitly stated and external to speaker's personal point of view.

Examples: Gives interpretation or explanation of some phenomena without implying that it is the only "correct" way of looking at it; elaborates or gives examples of a concept; points out relationships between examples and concepts or between two or more concepts; points out discrepancies between concept and examples; proposes hypothetical example or case to illustrate a point or raise a question.

10. Shows Positive Feeling

An act which implies positive evaluation of some behavior or interaction in the observational field, regardless of whether the referent is the self or some other person.

Examples: Any friendly act or overture, such as greeting or responding to a greeting; praising, approving, encouraging, rewarding, or showing active attention to other; sharing or sympathizing with other; expressions of satisfaction, enjoyment, or relief; joking or laughing "with" other.

11. Inhibits Communication

An act which implies unwillingness or inability to engage in the ongoing process of communication, regardless of whether the act stems from negative evaluation, internal tension, or disinterest.

Examples: Does not respond when response would ordinarily be expected; is cool, aloof, or disinterested in what is going on; is inattentive to or ignores a question or request; does not comply with a request; shows tension by blocking, "fright," etc.; accepts criticism or rebuff without reply.

12. Shows Negative Feeling

An act which implies active negative evaluation of some behavior or interaction in the observational field, regardless of whether the referrent is the self or some other person.

Examples: Disapproving, disparaging, threatening, discouraging other's behavior; lowering other's status; defending or asserting self; poking fun, belittling, or laughing "at" other; expressing fear, rage, hostility, disappointment, discouragement, displeasure, unhappiness, etc.

13. No Communication

The behavior occurring in the classroom is not relevant to teacher-pupil communication during a ten second interval.

14. Perfunctory Agreement or Disagreement

An act which implies indifference, usually made in a mechanical fashion.

II. FLANDERS MODEL

1.*	ACCEPTS FEELING; accepts and clarifies the feeling tone of the students in a nonthreatening manner. Feelings may be positive or negative. Predicting or recalling feelings are included.	Indirect Influence	TEACHER TALK
2.*	PRAISES OR ENCOURAGES: praises or encourages, student action or behavior. Jokes that release tension, not at the expense of another individual, nodding head, or saying, "um hm? " or "go on" are included.		
3.*	ACCEPTS OR USES IDEAS OF STUDENT: clarifying, building, or developing ideas suggested by a student. As teacher brings more of his own ideas into play, shift to category five.		
4.*	ASKS QUESTIONS: asking a question about content or procedure with the intent that a student answer.		
5.*	LECTURING: giving facts or opinions about content or procedure; expressing his own ideas, asking rhetorical questions.	Direct Influence	
6.*	GIVING DIRECTIONS: directions, commands, or orders to which a student is expected to comply.		
7.*	CRITICIZING OR JUSTIFYING AUTHORITY: statements intended to change student behavior from nonacceptable to acceptable pattern; bawling someone out; stating why the teacher is doing what he is doing; extreme self-reference.		
8.*	STUDENT TALK--RESPONSE: talk by students in response to teacher. Teacher initiates the contact or solicits student statement.		STUDENT TALK
9.*	STUDENT TALK--INITIATION: talk by students which they initiate. If "calling on" student is only to indicate who may talk next, observer must decide whether student wanted to talk. If he did, use this category.		
10.*	SILENCE OR CONFUSION: pauses, short periods of silence and periods of confusion in which communication cannot be understood by the observer.		

* There is NO scale implied by these numbers. Each number is classificatory, it designates a particular kind of communication event. To write these numbers down during observation is to enumerate, not to judge a position on a scale.

APPENDIX B

TABLES OF ADJUSTED MEANS

PRECEDING PAGE BLANK-NOT FILMED

I. CALIFORNIA TEST OF PERSONALITY

TABLE 103

California Test of Personality: Personal Adjustment
Adjusted Means by Sex, Grade, and Organization

Academic Year 1961-62

Sex	Grade	Washington Team	Washington Self-Contained	Longfellow Self-Contained
Girls	1	10.244	21.997	12.903
	2	20.482	20.993	19.401
	3	19.188	20.094	33.393
	4	19.068	19.668	22.521
	5	24.092	19.145	24.665
	6	23.446	28.357	23.959
Boys	1	16.904	27.515	23.117
	2	27.314	21.655	23.535
	3	14.708	25.236	28.321
	4	18.376	17.690	20.625
	5	17.828	21.107	23.567
	6	20.994	15.311	23.629

TABLE 104

California Test of Personality: Social Adjustment Adjusted Means by Sex, Grade, and Organization

Academic Year 1961-62

Sex	Grade	Washington Team	Washington Self-Contained	Longfellow Self-Contained
Girls	1	17.013	31.154	17.293
	2	27.349	21.466	23.815
	3	30.353	25.428	32.002
	4	18.819	29.922	28.758
	5	26.687	26.968	29.463
	6	22.951	31.112	29.841
Boys	1	25.947	30.834	22.161
	2	25.391	19.706	23.567
	3	26.641	27.298	27.856
	4	19.843	21.792	20.600
	5	21.897	26.068	26.459
	6	21.773	20.180	23.249

TABLE 105

California Test of Personality: Total Adjustment Adjusted Means by Sex, Grade, and Organization Academic Year 1961-62

Sex	Grade	Washington Team	Washington Self-Contained	Longfellow Self-Contained
Girls	1	16.436	43.581	16.891
	2	36.560	31.599	32.479
	3	38.060	34.869	55.516
	4	27.980	38.551	39.876
	5	40.720	34.385	43.509
	6	36.228	48.915	43.473
Boys	1	32.538	49.661	34.813
	2	43.730	30.315	36.721
	3	31.294	42.817	43.252
	4	28.470	30.659	30.976
	5	28.038	38.721	39.183
	6	31.710	24.775	36.827

II. CALIFORNIA ACHIEVEMENT TEST

TABLE 106

California Achievement Test: Reading Adjusted Means by Grade and Organization

Academic Year 1961-62

Grade	Washington Team	Washington Self-Contained	Longfellow Self-Contained
2	1.254	-14.300	23.154
3	3.026	9.989	69.576
4	-14.492	6.946	-23.637
5	9.777	54.561	18.739
6	-27.607	43.762	20.887

TABLE 107

California Achievement Test: Arithmetic Adjusted Means by Grade and Organization

Academic Year 1961-62

Grade	Washington Team	Washington Self-Contained	Longfellow Self-Contained
2	5.305	-6.189	-31.802
3	2.305	13.944	90.233
4	1.262	21.151	-25.599
5	12.090	101.644	28.887
6	-39.079	167.351	127.987

TABLE 108

California Achievement Test: Language Adjusted Means by Grade and Organization

Academic Year 1961-62

Grade	Washington Team	Washington Self-Contained	Longfellow Self-Contained
2	39.842	12.365	21.202
3	30.161	28.296	109.275
4	-14.325	26.751	-15.653
5	45.657	102.146	23.496
6	-11.026	59.589	-8.738

TABLE 109

California Achievement Test: Total Achievement Adjusted Means by Grade and Organization

Academic Year 1961-62

Grade	Washington Team	Washington Self-Contained	Longfellow Self-Contained
2	54.934	32.993	39.705
3	43.899	45.312	117.282
4	-6.144	28.005	-25.306
5	13.659	72.004	8.425
6	-46.895	64.218	15.025

PRECEDING PAGE BLANK-NOT FILMED

APPENDIX C

TEACHER AWARENESS

PRECEDING PAGE BLANK-NOT FILMED

Teacher Awareness Questionnaire (and Instructions)

Verbal instructions for the administration of the pupil information questionnaire.

(Read the following instructions to the teachers.)

This is part of the team-teaching project. Each of you will receive a questionnaire for fifteen pupils in your class or team. Please fill out as accurately as possible each of the eight questions for each child whose name appears on the top.

The standard that will be used to evaluate your answers will be the cumulative record.

We understand that the team is larger than the self-contained classes, and that the interns are at a greater disadvantage than the rest of the team members, but do the best you can.

(Pass out the questionnaires.)

Teacher ____________________
School ____________________
Grade ____________________

Child's Name ______________________

1. Age_______ yrs. _______ mos.

2. Number of siblings living with family _______.

3. Is the child living with both natural parents? Yes No (circle one)

4. Father's (or step-father's) occupation (be as specific as you can):

   ____

5. Mother's (or Step-mother's) occupation (be as specific as you can):

   ____

6. IQ (or reading readiness score for first grade only) as recorded in the cumulative record: _____________.

7. Overall grade placement score (CAT) to the nearest tenth of a grade as recorded in the cumulative record: ______________.

8. Fourth, Fifth and Sixth Grade Teachers Only

   Given a scale from one to three, circle to the nearest tenth of a point what you feel the child's sociometric rating by the rest of the class would be:

   1.0 1.1 1.2 1.3 1.4 1.5 1.6 1.7 1.8 1.9 2.0

   2.1 2.2 2.3 2.4 2.5 2.6 2.7 2.8 2.9 3.0

9. Are there any comments you would like to make? E. g., have you answered any of the questions with more recent data than may appear in the cumulative record?

Please examine one of the questionnaires. Some of the questions need qualification.

Question number one is straight-forward.

On question number two, enter the number of siblings that appears on the latest Entrance Questionnaire for the pupil. This includes step and foster siblings.

Are there any questions on items number three, four, or five?

For item number six, first grade teachers record the percentile rank of the pupil on the latest Reading Readiness Test.

All others record the latest intelligence measure that appears in the cumulative record as of March 30. DO NOT RECORD A PERCENTILE RANK OR MENTAL AGE, BUT AN IQ MEASURE.

For question number seven, disregard the reference to the cumulative record. Instead, estimate the overall grade placement score that you believe the pupil obtained on the administration of the California Achievement Test last October. Please remember that Madison norms are higher than national norms. Thus, if you considered a third grader as average in October, you would not record 3. 1, but in comparison to national norms, you would add three-tenths of a grade (3. 4). A rule of thumb is to add .3 to all scores to match Madison scores to the national norms. Are there any questions?

On item number eight, every member of the class was rated either 1, 2, or 3 by each member of the rest of the class in a separate sociometric study last fall. One is "Do not choose to play with. ": Two is "Sometimes like to play with. ": Three is "Especially like to play with. " Then a class average was computed for each child. Remember, the lower the sociometric rating the student receives from the class, the lower the average.

Are there any questions?

In case you have answered a question with other information than appears in the cumulative record, please make a note of it under item number nine.

Please feel free to comment.

Remember, be sure to answer all questions. You must, of course, work alone.

You may leave when you're finished. Thank you.

BIBLIOGRAPHY

PRECEDING PAGE BLANK-NOT FILMED

1. Ackerlund, G. "Some Teacher Views on the Self-Contained Classroom." Phi Delta Kappan, April 1959, 40, 283-285.

2. Adams, A. S. "Operation Co-Teaching, Dateline: Oceano, California." The Elementary School Journal, January 1962, 62, 203-212.

3. Administrator's Guide to Team Teaching. Educational Executives' Overview, April 1963, 4, 54-55.

4. Anderson, Harold H., and Brewer, Joseph E. "Studies of Teachers' Classroom Personalities, II: Effects of Teachers' Dominative and Integrative Contacts on Children's Behavior." Applied Psychology Monographs, No. 8, June 1946.

5. Anderson, R. H. "Team Teaching." National Education Association Journal, March 1961, 50, 52-54.

6. Anderson, R. H. "Team Teaching." The Education Digest, May 1961, 26, 5-7.

7. Anderson, R. H. "Team Teaching in Action." The Nation's Schools, May 1960, 65, 62+.

8. Anderson, R. H. "Team Teaching in the Elementary School." The Education Digest, November 1959, 25, 26-28.

9. Anderson, R. H. "Time Allotment: A View on Priorities." Educational Horizons, Winter 1958, 37, 34-39.

10. Anderson, R. H., Hagstrom, E. A., and Robinson, W. A. "Team Teaching in an Elementary School." School Review, Spring 1960, 68, 71-84.

11. Anderson, R. H. and Mitchell, D. P. "Team Teaching, New Learning Concepts Demand Changes in School Plant Design." The Nation's Schools, June 1960, 65, 75+.

12. Andree, R. G. "Large Classes and Effective Teaching." The Clearing House, February 1959, 33, 334-336.

13. Bales, Robert F. Interaction Process Analysis. Cambridge, Mass.: Addison-Wesley, 1953.

14. Bank Street College of Education, New York. "The Classroom Processes Study." Working paper prepared for the Work Conference held by the staffs of four NIMH-related projects at the University of Wisconsin, November 10-13, 1960.

15. Baynham, D. "Selected Staff Utilization Projects in California, Georgia, Colorado, Illinois, Michigan, and New York." National Association of Secondary School Principals Bulletin, January 1962, 46, 14-98.

16. Becker, H. A. "Team Teaching." Instructor, June 1962, 71: 43-45.

17. Bissex, H. S. "Newton Plan Challenges Traditions of Class Size." The Nation's Schools, March 1960, 65, 60-64.

18. Bleifeld, M. "What Responsibilities for the Principal in Organizing, Supervising, and Evaluating Teaching Teams?" National Association of Secondary School Principals Bulletin, April 1961, 45, 117-118.

19. Bloomenshine, L. L. "Team Teaching in San Diego - The First Year." National Association of Secondary School Principals Bulletin, January 1960, 44, 181-196.

20. Bloomenshine, L. L. and Brown, T. M. "San Diego, California, Conducts Two-Year Experiment with Team Teaching." National Association of Secondary School Principals Bulletin, January 1961, 45, 147-166.

21. Boutwell, W. D. "What's Happening in Education?" The PTA Magazine, May 1963, 57, 25-26.

22. Broadhead, F. C. "Pupil Adjustment in the Semi-Departmental Elementary School." Elementary School Journal, April 1960, 60, 385-390.

23. Brookover, Wilbur B. "Person-Person Interaction between Teachers and Pupils, and Teaching Effectiveness." Journal of Educational Research, December 1940, 34, 272-87.

24. Brown, B. F. "A New Design for Individual Learning." Journal of Secondary Education, October 1962, 37, 368-375.

25. Brownell, J. A. and Taylor, H. A. "Theoretical Perspectives for Teaching Teams." Phi Delta Kappan, January 1962, 43, 150-157.

26. Bruntz, G. G. "The Team Approach to Social Science Teaching." The High School Journal, April 1960, 43, 370-374.

27. Bush, R. N. "A Searching Appraisal of New Developments." Journal of Secondary Education, October, 1962, 37, 321-326.

28. Bush, R. N. "The Team Teaching Bandwagon." California Journal of Secondary Education, March 1960, 35, 207-208.

29. California Short Form Test of Mental Maturity. Devised by Elizabeth P. Sullivan, Willis W. Clark, and Ernest W. Tiegs. California Test Bureau, Monterey, California, 1957.

30. Cashen, V. M. "Using Specialists as a Team." Educational Leadership, November 1961, 19, 115-117.

31. Clark, A. B. "Administering Team Teaching in East Side District, San Jose, California." National Association of Secondary School Principals Bulletin, January 1962, 46, 141-144.

32. Clark, A. B. "An Appraisal of Team Teaching." Journal of Secondary Education, November 1961, 36, 441-444.

33. Clawson, H. A. "Science Lecture and Team Approaches in English are Tried in Mattoon High School." National Association of Secondary School Principals Bulletin, January 1959, 43, 245-247.

34. Cobett, E. M. "A Different Approach to Team Teaching." Ohio Schools, November 1960, 38, 10-11.

35. Cooper, W. L. "J. Sterling Morton High School and Junior College, Cicero, Illinois, Uses Tapes, Language Laboratories, and Team Teaching." National Association of Secondary School Principals Bulletin, January 1961, 45, 79-84.

36. Cornell, Francis G., Lindvall, Carl M., and Saupe, Joe L. "An Exploratory Measurement of Individualities of Schools and Classrooms." University of Illinois Bulletin, June 1953.

37. Crandell, E. and Piel, W. "Birmingham Tries Team Teaching Experiment." Michigan Education Journal, January 1961, 38, 344-345+.

38. "Critical Look at Team Teaching." Instructor, October 1961, 71, 39-42.

39. Cunningham, L. L. "Keys to Team Teaching." Overview, October 1960, 1, 54-55.

40. Cunningham, L. L. "When Is a Team a Team?" The High School Journal, October 1961, 45, 7-13.

41. Dagnall, W. "Team Teaching." The Times Educational Supplement, May 1962, 2453, 1080.

42. Dean, R. B. "Team Teaching in the Elementary Schools." The American School Board Journal, December 1962, 145, 5-6.

43. Dean, S. E. "Team Teaching: A Review." School Life, September 1961, 44, 5-8.

44. Dillion, C. L. "Taylorville, Illinois, Senior High School Uses Tape Recorders, Team Teaching, and Large-Group Instruction to Improve Staff Utilization." National Association of Secondary School Principals Bulletin, January 1961, 45, 178-188.

45. Drummond, H. D. "Team Teaching: An Assessment." Educational Leadership, December 1961, 19, 160-165.

46. Edwards, Allen L. Experimental Design in Psychological Research, New York: Holt, Rinehart, and Winston, 1960.

47. Elliott, R. W. "Team Teaching: Effective In-Service Training." The American School Board Journal, February 1962, 144, 19.

48. Empey, D. W. "Student Self-Direction, Flexible Scheduling, and Team Teaching." National Association of Secondary School Principals Bulletin, February 1963, 47, 118-124.

49. Evans, H. and Womer, G. S. "Stirrings in the Big Cities: Cleveland." National Education Association Journal, November 1962, 51, 50-53.

50. Ewan, S. N. "Class Size." National Association of Secondary School Principals Bulletin, March 1932, 16, 124-129.

51. Falk, P. H. "The Improvement of Instruction." The Bulletin, publication of the Madison Education Association, Madison, Wisconsin, April 1960, 30, 1.

52. Fischler, A. S. "The Use of Team Teaching in the Elementary School." School Science and Mathematics, April 1962, 62, 281-288.

53. Fischler, A. S., and Shoresman, P. B. "Team Teaching in the Elementary School: Implications for Research in Science Instruction." Science Education, December 1962, 46, 405-415.

54. Fisher, M. O. "Team Teaching in Houston." English Journal, December 1962, 51, 628-631.

55. Flanders, Ned A. Interaction Analysis in the Classroom. Minneapolis, Minn.: College of Education, University of Minnesota, August 1960.

56. Foster, V. E. "Teaching by Teams." International Journal of Religious Education, March 1957, 33, 18-19.

57. Gilchrist, R. S. "Promising Practices in Education, II." Phi Delta Kappan, March 1960, 41, 269-274.

58. Gilchrist, R. S.; Bergstrom, H. E.; Duval, F. H., and Mueller, E. H. "Curriculum News." Educational Leadership, May 1960, 17, 517-520.

59. Giltinan, B. "We Solved the Problem of Size." The English Journal, February 1963, 52, 89-93.

60. Glancy, P. B. "Brookside Junior High, Sarasota, Florida, Stives for Quality Education." National Association of Secondary School Principal Bulletin, January 1962, 157-160.

61. Goodlad, J. I. "News and Comment." The Elementary School Journal, October 1958, 59, 1-17.

62. Goodlad, J. I. "Room to Live and Learn." Childhood Education, April 1954, 30, 355-361.

63. Greenshields, M. J. "Big Timber's Teacher-Helper Plan." American School Board Journal, April 1942, 104, 20.

64. Gross, C. E. "Team Teaching in Pittsburgh." Education Digest, November 1962, 28, 12-15.

65. Gross, R. E. "Emerging Horizon for the Social Studies." Social Education, January 1960, 24, 21-24.

66. Haan, C. T., and Adams, R. C. "Experimentation at Fremont High School." Journal of Secondary Education, May 1962, 37, 274-279.

67. Hagstrom, E. A. "New Opportunity for Outstanding Teachers." Grade Teacher, January 1961, 78, 13+.

68. Hamalainen, A. E. "Some Current Proposals and Their Meaning." Educational Leadership, February 1959, 16, 271-274+.

69. Hanvey, R., and Tenenberg, M. S. "University of Chicago Laboratory School, Chicago, Illinois, Evaluates Team Teaching." National Association of Secondary School Principals Bulletin, January 1961, 45, 189-197.

70. Harrison, W. J. "Team Teaching at Muskegon, Michigan, Senior High School." National Association of Secondary School Principals Bulletin, January 1962, 46, 238-242.

71. Heathers, G. "The Dual Progress Plan." Educational Leadership, November 1960, 18, 89-91.

72. Heller, M. P., and Belford, E. "Team Teaching and Staff Utilization in Ridgewood High School." National Association of Secondary School Principals Bulletin, January 1962, 46, 105-122.

73. Hemeyer, W., and McGrew, J. B. "Big Ideas for Big Classes." School Review, Autumn 1960, 68, 308-317.

74. Holmes, D., and Harvey, L. F. "An Evaluation of Two Methods of Grouping." Educational Research Bulletin, November 1956, 35, 213-222.

75. Hoppock, A. "Team Teaching: Form without Substance." The Education Digest, September 1961, 27, 9-11.

76. Horn, E. G. "Team Approach Adds Flexibility to Our First-Grade Program." Illinois Education, November 1961, 50, 114-115.

77. Hosley, Charles T. Learning Outcomes of 6th Grade Pupils under Alternate Grade Organization Patterns. Unpublished Ph.D. thesis, Stanford University, 1954.

78. Howe, H. "The Curriculum, the Team, and the School: An Examination of Relationships." Journal of Secondary Education, October 1962, 37, 353-361.

79. Hull, J. H. "Multigrade Teaching." The Nation's Schools, July 1958, 62, 33-36.

80. Iannizzi, E., and Robinson, C. "Team Teaching, Should You Attempt It?" The Journal of Business Education, April 1963, 38, 274-276.

81. Iowa Test of Basic Skills, Manual for Administrators, Supervisors, and Counselors; Houghton-Mifflin Company, Boston, Mass., 1956.

82. Iowa Test of Basic Skills, Teachers Manual; Houghton-Mifflin Company, Boston, Mass., 1956.

83. Jensen, L., Riggle, W. B., Merkley, P., Nielsen, V., and Rudy, R. "Eighth Grade Team Teaching at the Roosevelt Junior High School." California Journal of Secondary Education, April 1960, 35, 236-243.

84. Johnson, R. H. and Lobb, M. D. "Jefferson County, Colorado, Completes Three-Year Study of Staffing, Changing Class Size, Programming, and Scheduling." National Association of Secondary School Principals Bulletin, January 1961, 45, 57-78.

85. Johnson, R. H., Lobb, M. D., and Patterson, G. "Continued Study of Class Size, Team Teaching, and Scheduling in Eight High Schools in Jefferson County, Colorado." *National Association of Secondary School Principals Bulletin*, January 1959, 43, 99-103.

86. Johnson, R. H., Lobb, M. D., and Swenson, L. G. "An Extensive Study of Team Teaching and Schedule Modification in Jefferson County, Colorado, School District R-1." *National Association of Secondary School Principals Bulletin*, January 1960, 44, 79-93.

87. Jonsson, S. R. "Team Teaching? Enthusiasm Is High." *New York State Education*, November 1962, 50, 14-16.

88. Katz, L. and Proctor, C. H. "The Concept of Configuration of Interpersonal Relations in a Group as a Time-dependent Stochastic Process." *Psychometrika*, December, 1959, 24, 317-327.

89. Keach, S. J. "Team Teaching Is Exciting!" *International Journal of Religious Education*, September 1957, 34, 12-13.

90. Keliher, A. V. "Team Teaching." *Highpoints*, May 1962, 44, 65-68.

91. Keppel, F., Shaplin, J. T., and Robinson, W. M. "Recent Developments at the Harvard Graduate School of Education." *The High School Journal*, February 1960, 43, 242-261.

92. Kidd, J. W. "The Question of Class Size." *The Journal of Higher Education*, November 1959, 23, 440-444.

93. King, Jr., A. R. "Planning for Team Teaching: The Human Considerations." *The Journal of Secondary Education*, October 1962, 37, 363-367.

94. Kopp, O. W. "Grouping Pupils in the Elementary School." *New York State Education*, May 1957, 44, 540-541+.

95. Kramer, C. Y. "Extension of Multiple Range Tests to Group Means with Unequal Numbers of Replications." *Biometrics*, September, 1956, 12, 307-310.

96. Lambert, P. "Team Teaching for the Elementary School." *Educational Leadership*, November 1960, 18, 85-88+.

97. Lambert, P. "Team Teaching for Today's World." *Teachers College Record*, March 1963, 64, 480-486.

98. Lambert, P., and Boyan, N. J. "Team Teaching - Is Your School Ready for It?" *Michigan Journal of Secondary Education*, Spring 1961, 2, 3.

99. Larmee, R. A., and Ohm, R. "University of Chicago Laboratory School Freshman Project Involves Team Teaching, New Faculty Position, and Regrouping of Students." National Association of Secondary School Principals Bulletin, January 1960, 44, 275-289.

100. Leonard, G. B. "The Truth about the Teacher Crisis." Look, February, 1956, 20, 4.

101. Lewis, W. W., Withall, John, and Newell, John M. A Description of Two Instructional Approaches. University of Wisconsin Mental Health - Teacher Education Research Project, Madison, Wis., November 1960, mimeo.

102. Lobb, M. D., Noall, M. F., and Slichenmyer, H. L. "What Are Some Promising Practices in Team Teaching?" National Association of Secondary School Principals Bulletin, April, 1960, 44, 2-7.

103. Lord, J. C. E. "Team Teaching Should Be Tailored to the Individual School Situation." Business Education World, April 1963, 43, 10-11+.

104. Loretan, J. O. "Team Teaching: Plus and Minus in New York City's Junior High Schools." National Association of Secondary School Principals Bulletin, January 1962, 46, 135-140.

105. Luchins, Abraham S., and Luchins, Edith H. "Children's Attitudes towards Homogeneous Groupings." Pedagog Seminar, March 1948, 72, 3-9.

106. Mahoney, W. M. "Try Co-ordinate Teaching." The American School Board Journal, November 1959, 139, 13-14.

107. Marsh, R. "Team Teaching - New Concept?" The Clearing House, April 1961, 35, 496-499.

108. Mayer, F. C., and Woolridge, J. H. "Preparing for Team Teaching at West Clermont." The American School Board Journal, July 1962, 145, 10.

109. McCallum, T. E., Henry, C. D., and Nesbitt, W. O. "Snyder, Texas, Continues Team Teaching." National Association of Secondary School Principals Bulletin, January 1961, 45, 261-265.

110. McCurties, H. L., and Schycker, B. H. "Two Teachers and A One-Room 'School'." Chicago Schools Journal, December 1962, 44, 131-133.

111. Mental Health-Teacher Education Research Project. National Institute of Mental Health Project # 2M-6624.

112. Miel, A. "The Self-Contained Classroom: An Assessment." Teachers College Record, February 1958, 59, 282-291.

113. Mitchell, W. B. "Why Try Team Teaching?" National Association of Secondary-School Principals Bulletin, January 1962, 46, 247-252.

114. Montague, D. O. "Team Teaching in Berkeley's Burbank Junior High School." Journal of Secondary Education, November 1961, 36, 420-422.

115. Nagle, A. J. "The Team Concept in Organizing a Secondary-School Program." National Association of Secondary School Principals Bulletin, April 1959, 43, 86-87.

116. Nelson, J., Hahn, R. O., and Robinson, G. A. "Team Teaching for Teacher Education: The New Approach." Journal of Teacher Education, September 1961, 12, 380-382.

117. Nesbitt, W. O. "Big Classes in Texas." Educational Screen and Audiovisual Guide, November 1959, 38, 594-596.

118. Noall, M. F., and Jensen, L. "Team Teaching at Roosevelt Junior High School, Duchesne County, Utah." National Association of Secondary School Principals Bulletin, January 1960, 44, 156-163.

119. Noall, M. F., Riggle, W., and Jensen, L. "Teacher - Team Project, Roosevelt Junior High School, Duchesne County School District, Utah." National Association of Secondary School Principals Bulletin, January 1961, 45, 234-238.

120. Noall, M. F. and Rose, G. "Team Teaching at the Wahlquist Junior High School, Weber County, Utah." National Association of Secondary School Principals Bulletin, January 1960, 44, 164-171.

121. Norton, Monte S. "Approaches to Team Teaching." National Association of Secondary School Principals Bulletin, October 1960, 44, 89-92.

122. Park, C. B. "The Bay City Experiment... As Seen by the Director." Journal of Teacher Education, June 1956, 7, 101-110.

123. Park, C. B. "The Teacher-Aide Plan." The Nation's Schools, July 1955, 56, 45-55.

124. Partridge, A. R. "Staff Utilization in Senior High School." Educational Leadership, January 1961, 18, 217-221.

125. Perkins, B., et al. "Teamwork Produces Audio-Visual Techniques." Grade Teacher, June 1960, 77, 55-72.

126. Pistor, F. A. "Evaluating Newer School Practices by the Observational Method," in Appraising the Elementary-School Program. 16th Yearbook. Department of Elementary Principals, National Education Association, 1937, 377-89.

127. Pitruzzello, P. R. "A Report on Team Teaching." The Clearing House, February 1962, 36, 333-336.

128. Ploghoft, M. E. "Another Look at Team Teaching." The Clearing House, December 1961, 36, 219-221.

129. Polos, N. C. "The Teaching Team in Action." Journal of Secondary Education, November 1961, 36, 415-419.

130. Rehwoldt, W., and Hamilton, W. "Why Group by Grade Level." Grade Teacher, January 1959, 75, 18-19+.

131. Rzepka, L. "Team Teaching in the Elementary School. What Is It?" Ohio Schools, January 1962, 40, 14-15+.

132. Siegel, Sidney. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill Book Company, Inc., 1956.

133. Singer, I. J. "Summer Meeting stresses Technology Theme." Audio-visual Instruction, September 1961, 6, 352-3.

134. Singer, I. J. "Survey of Staff Utilization Practices in Six States." National Association of Secondary School Principals Bulletin, January 1962, 46, 1-13.

135. Smith, D. V. "Experiments in Handling Larger Classes." The English Journal, May 1931, 20, 371-378.

136. Smith, G. "Experimentation at Verdugo Hills High School." Journal of Secondary Education, November 1961, 36, 433-440.

137. Smith, V. H. "Team Teaching Has Advantages." The English Journal, April 1960, 49, 242-244+.

138. Stetson, G. A., and Harrison, J. P. "A Junior High School Designed for Team Teaching." The American School Board Journal, May 1960, 140, 38-42.

139. Stiles, L. J. "Individual and Team Teaching." Wisconsin Journal of Education, June 1960, 92, 7-10+.

140. Stoddard, G. D. in Hunnicutt, C. W., ed., Education 2000 AD, Syracuse University Press, 1956, 141-164.

141. Sweet, R., and Dunn-Rankin, P. "An Experiment in Team Teaching Seventh Grade Arithmetic." School Science and Mathematics, May 1962, 62, 341-44.

142. Szabo, L. J. "Team Teaching: Honors Student Undergo Experiment." New York State Education, October 1961, 49, 12-13.

143. Tabachnick, B. R. "Possibilities for Team Teaching." Wisconsin Journal of Education, November 1962, 95, 5-7.

144. Taylor, H. A. "Claremont Graduate School Program for Team Teaching." The High School Journal, February 1960, 43, 277-282.

145. "Team Teaching Proves Worthwhile in Chicago." Chicago Schools Journal, March 1963, 44, 283-284.

146. Thomson, S. D. "Can Team Teaching Aid Learning?" Journal of Secondary Education, November 1961, 36, 423-429.

147. Thorp, Louis P., Clark, Willis W., and Tiegs, Ernest W. California Test of Personality Manual (all levels). California Test Bureau, Monterey, California, Revised 1953.

148. Tiegs, Ernest W. and Clark, Willis W. California Achievement Tests Complete Battery Manual (Separate Manuals for Lower Primary, Upper Primary, Elementary, and Junior High levels). California Test Bureau, Los Angeles, California, 1957 edition.

149. Time, Talent, and Teachers. The Ford Foundation, New York, N. Y., 1960.

150. Tink, A. K. A summary of the presentation made by A. K. Tink at the Forty-fifth Annual Convention of the National Association of Secondary School Principals at Cobo Hall, Detroit, Michigan, Feb., 1961 as reported in National Association of Secondary School Principals Bulletin, April, 1961, 45, 119-120.

151. Trump, J. Lloyd. "A Look Ahead in Secondary Education." National Association of Secondary School Principals Bulletin, January 1958, 42, 5-15.

152. Varner, G. F. "Team Teaching in Johnson High School, St. Paul, Minnesota." National Association of Secondary School Principals Bulletin, January 1962, 46, 161-166.

153. Warburton, J. T. "An Experiment in Large Group Instruction." Journal of Secondary Education, November 1961, 36, 430-432.

154. Ward, J. O. "Another Plan for Co-ordinate Teaching." The American School Board Journal, February 1960, 140, 10.

155. Weiss, T. M., and Morris, M. S. "A Critique of the 'Team Approach'." The Educational Forum, January 1960, 24, 207-208.

156. "What Size Elementary Classes?" The Nation's Schools, April 1945, 35, 42.

157. "What Size High School Classes?" The Nation's Schools, June 1945, 35, 29.

158. Withall, John G. "Development of a Technique for the Measurement of Social-Emotional Climate in Classrooms." Journal of Experimental Education, March 1949, 17, 347-61.

159. Wood, M. "Team Teaching Can Help in Smaller Classes, Too." Business Education World, October 1962, 43, 30-31.

160. Wrightstone, J. Wayne. Appraisal of Newer Elementary School Practices. New York, N.Y.: Teachers College, 1938.